



Europäisches Patentamt
European Patent Office
Office européen des brevets



⑪ Publication number:

0 346 710 B1

⑫

EUROPEAN PATENT SPECIFICATION

⑯ Date of publication of patent specification: 10.11.93 ⑯ Int. Cl.5: **C12N 15/12, C12N 5/10, C12P 21/02**
⑯ Application number: 89110096.8
⑯ Date of filing: 03.06.89

⑯ cDNAs coding for members of the carcinoembryonic antigen family.

⑯ Priority: 16.06.88 US 207678
21.11.88 US 274107

⑯ Date of publication of application:
20.12.89 Bulletin 89/51

⑯ Publication of the grant of the patent:
10.11.93 Bulletin 93/45

⑯ Designated Contracting States:
AT BE CH DE ES FR GB GR IT LI NL SE

⑯ References cited:
EP-A- 263 933
EP-A- 0 212 880

BIOCHEM. BIOPHYS. RES. COMMUN., vol. 142, no. 2, 30th January 1987, pages 511-518;
R. OIKAWA et al.: "Primary structure of human carcinoembryonic antigen (CEA) deduced from cDNA sequence"

MOL. CELL. BIOL., vol. 7, 1987, page 3221-3230; R. BEAUCHEMIN et al.: "Isolation and characterization of full-length functional cDNA clones for human carcinoembryonic antigen"

⑯ Proprietor: MILES INC.
One Mellon Center
500 Grant Str.
Pittsburgh, PA 15219-2502(US)

⑯ Inventor: Barnett, Thomas R., Dr.
27 Jeffrey Road
East Haven, CT 06513(US)
Inventor: Elting, James J., Dr.
5 Heatherwood Drive
Madison, CT 06443(US)
Inventor: Kamarck, Michael E.
86 Russell Road
Bethany, CT 06525(US)
Inventor: Kretschmer, Axel, Dr.
Richard-Zörner-Strasse 32
D-5060 Bergisch Gladbach 1(DE)

⑯ Representative: Dänner, Klaus, Dr. et al
Bayer AG
Konzernverwaltung RP
Patente Konzern
D-51368 Leverkusen (DE)

EP 0 346 710 B1

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid (Art. 99(1) European patent convention).

PROC. NATL. ACAD. SCI. USA, vol. 85, September 1988, pages 6959-6963; Y. HINODA et al.: "Molecular cloning of a cDNA coding biliary glycoprotein I: primary structure of a glycoprotein immunologically crossreactive with carcinoembryonic antigen"

GENE, vol. 71, no. 2, November 1988, pages 439-449; B.C. ROONEY et al.: "Molecular cloning of a cDNA for human pregnancy-specific B1-glycoprotein: homology with human carcinoembryonic antigen and related proteins"

Description**BACKGROUND OF THE INVENTION****5 Field of the Invention**

The present invention concerns nucleic acid sequences which code for carcinoembryonic antigen (CEA) antigen family peptide sequences.

10 Background Information

15 Carcinoembryonic antigen was first described by Gold and Freedman, *J. Exp. Med.*, 121, 439-462, (1965). CEA is characterized as a glycoprotein of approximately 200,000 molecular weight with 50-60% by weight of carbohydrate. CEA is present during normal human fetal development, but only in very low concentration in the normal adult intestinal tract. It is produced and secreted by a number of different tumors.

20 CEA is a clinically useful tumor marker for the management of colorectal cancer patients. CEA can be measured using sensitive immunoassay methods. When presurgical serum levels of CEA are elevated, a postsurgical drop in serum CEA to the normal range typically indicates successful resection of the tumor. Postsurgical CEA levels that do not return to normal often indicate incomplete resection of the tumor or the presence of additional tumor sites in the patient. After returning to normal levels, subsequent rapid rises in serum CEA levels usually indicate the presence of metastases. Slower postsurgical rises from the normal level are most often interpreted to indicate the presence of new primary tumors not previously detected. Post surgical management of colon cancer patients is thus facilitated by the measurement of CEA.

25 CEA is a member of an antigen family. Because of this, the immunoassay of CEA by presently available methods is complicated by the fact that CEA is but one of several potentially reactive antigens. There have been at least sixteen CEA-like antigens described in the literature. Since some of these appear to be the same antigen described by different investigators, the actual number of different antigens is somewhat less than this number. Nonetheless, there is a complex array of cross-reactive antigens which 30 can potentially interfere with an immunoassay of the CEA released by tumors. It is known that serum levels of CEA-like antigens are elevated in many non-cancerous conditions such as inflammatory liver diseases and also in smokers. It is important that immunoassays used for the monitoring of cancer patient status not be interfered with by these other CEA-like antigens. Conversely, it is important to be able to distinguish the 35 antigens by immunoassays because of the possibility that different tumor types may preferentially express different forms of CEA. If so, then the ability to reliably measure the different forms of CEA can provide the means to diagnose or more successfully treat different forms of cancer.

The members of the "CEA family" share some antigenic determinants. These common epitopes are not useful in distinguishing the members of the antigen family and antibodies recognizing them are of little use for measuring tumor-specific CEA levels.

40 U.S.P. 3,663,684, entitled "Carcinoembryonic Antigen and Diagnostic Method Using Radioactive Iodine", concerns purification and radioiodination of CEA for use in a RIA.

45 U.S.P. 3,697,638 describes that CEA is a mixture of antigens (components A and B in this case). U.S.P. 3,697,638 mentions methods for separating and radioiodinating each component and their use in specific RIA's.

50 U.S.P. 3,852,415, entitled "Compositions for Use in Radioimmunoassay, as Substitute for Blood Plasma Extract in Determination of Carcinoembryonic Antigen" relates to the use of a buffer containing EDTA and bovine serum albumin as a substitute for plasma as a diluent for CEA RIA's.

55 U.S.P. 3,867,363, entitled "Carcinoembryonic Antigens", is directed to the isolation of CEA components A and B, their labelling and use in a RIA.

U.S.P. 3,927,193, entitled "Localization of Tumors by Radiolabelled Antibodies", concerns the use of radiolabelled anti-CEA antibodies in whole body tumor imaging.

U.S.P. 3,956,258, entitled "Carcinoembryonic Antigens", relates to the isolation of CEA components A and B.

U.S.P. 4,086,217, entitled "Carcinoembryonic Antigens", is directed to the isolation of CEA components A and B.

U.S.P. 4,140,753, entitled "Diagnostic Method and Reagent", concerns the purification of a CEA isomer called CEA-S1 and its use in a RIA.

U.S.P. 4,145,336, entitled "Carcinoembryonic Antigen Isomer", relates to the antigen CEA-S1.

U.S.P. 4,180,499, entitled "Carcinoembryonic Antigens", describes a process for producing CEA component B.

U.S.P. 4,228,236, entitled "Process of Producing Carcinoembryonic Antigen", is directed to the use of the established cell lines LS-174T and LS-180 or clones or derivatives thereof for the production of CEA.

5 U.S.P. 4,272,504, entitled "Antibody Adsorbed Support Method for Carcinoembryonic Antigen Assay", concerns two concepts for the radioimmunoassay of CEA. First, U.S.P. 4,272,504 relates to a sample pretreatment in the form of heating to 65 to 85°C at pH 5 to precipitate and eliminate extraneous protein. Second, it describes the use of a solid phase antibody (either on beads or tubes) as a means to capture analyte and radiolabelled CEA tracer.

10 U.S.P. 4,299,815, entitled "Carcinoembryonic Antigen Determination", concerns diluting a CEA sample with water and pretreating by heating to a temperature below which precipitation of protein will occur. The pretreated sample is then immunoassayed using RIA, EIA, FIA or chemiluminescent immunoassay.

15 U.S.P. 4,349,528, entitled "Monoclonal Hybridoma Antibody Specific for High Molecular Weight Carcinoembryonic Antigen", is directed to a monoclonal antibody reacting with 180 kD CEA, but not with other molecular weight forms.

U.S.P. 4,467,031, entitled "Enzyme-Immunoassay for Carcinoembryonic Antigen", relates to a sandwich enzyme immunoassay for CEA in which the first of two anti-CEA monoclonal antibodies is attached to a solid phase and the second monoclonal is conjugated with peroxidase.

20 U.S.P. 4,489,167, entitled "Methods and Compositions for Cancer Detection", describes that CEA shares an antigenic determinant with alpha-acid glycoprotein (AG), which is a normal component of human serum. The method described therein concerns a solid-phase sandwich enzyme immunoassay using as one antibody an antibody recognizing AG and another antibody recognizing CEA, but not AG.

U.S.P. 4,578,349, entitled "Immunoassay for Carcinoembryonic Antigen (CEA)", is directed to the use of high salt containing buffers as diluents in CEA immunoassays.

25 EP 113072-A, entitled "Assaying Blood Sample for Carcinoembryonic Antigen - After Removal of Interfering Materials by Incubation with Silica Gel", relates to the removal from a serum of a plasma sample of interfering substances by pretreatment with silica gel. The precleared sample is then subjected to an immunoassay.

EP 102008-A, entitled "Cancer Diagnostics Carcinoembryonic Antigen - Produced from Perchloric Acid 30 Extracts Without Electrophoresis", relates to a procedure for the preparation of CEA from perchloric acid extracts, without the use of an electrophoresis step.

EP 92223-A, entitled "Determination of Carcinoembryonic Antigen in Cytosol or Tissue - for Therapy Control and Early Recognition of Regression", concerns an immunoassay of CEA, not in serum or plasma, but in the cytosol fraction of the tumor tissue itself.

35 EP 83103759.6, entitled "Cytosole-CEA-Measurement as Predictive Test in Carcinoma, Particularly Mammacarcinoma", is similar to EP 92223-A.

EP 83303759, entitled "Monoclonal Antibodies Specific to Carcinoembryonic Antigen", relates to the production of "CEA specific" monoclonal antibodies and their use in immunoassays.

WO 84/02983, entitled "Specific CEA-Family Antigens, Antibodies Specific Thereto and Their Methods 40 of Use", is directed to the use of monoclonal antibodies to CEA-meconium (MA)-, and NCA-specific epitopes in immunoassays designed to selectively measure each of these individual components in a sample.

All of the heretofore CEA assays utilize either monoclonal or polyclonal antibodies which are generated by immunizing animals with the intact antigen of choice. None of them address the idea of making 45 sequence specific antibodies for the detection of a unique primary sequence of the various antigens. They do not cover the use of any primary amino acid sequence for the production of antibodies to synthetic peptides or fragments of the natural product. They do not include the concept of using primary amino acid sequences to distinguish the CEA family members. None of them covers the use of DNA or RNA clones for isolating the genes with which to determine the primary sequence.

DEFINITIONSNucleic Acid Abbreviations

5	A	adenine
	G	guanine
	C	cytosine
	T	thymidine
10	U	uracil

Amino Acid Abbreviations:

15	Asp	aspartic acid
	Asn	asparagine
	Thr	threonine
	Ser	serine
20	Glu	glutamic acid
	Gln	glutamine
	Pro	proline
25	Gly	glycine
	Ala	alanine
	Cys	cysteine
30	Val	valine
	Met	methionine
	Ile	isoleucine
	Leu	leucine
35	Tyr	tyrosine
	Phe	phenylalanine
	Trp	tryptophan
40	Lys	lysine
	His	histidine
	Arg	arginine

45 Nucleotide - A monomeric unit of DNA or RNA containing a sugar moiety (pentose), a phosphate, and a nitrogenous heterocyclic base. The base is linked to the sugar moiety via the glycosidic carbon (1' carbon of the pentose) and that combination of base and sugar is called a nucleoside. The base characterizes the nucleotide. The four DNA bases are adenine ("A"), guanine ("G"), cytosine ("C"), and thymine ("T"). The four RNA bases are A, G, C and uracil ("U").

50 DNA Sequence - A linear array of nucleotides connected one to the other by phosphodiester bonds between the 3' and 5' carbons of adjacent pentoses.

55 Functional equivalents - It is well known in the art that in a DNA sequence some nucleotides can be replaced without having an influence on the sequence of the expression product. With respect to the peptide this term means that one or more amino acids which have no function in a particular use can be deleted or replaced by another one.

Codon - A DNA sequence of three nucleotides (a triplet) which encodes through mRNA an amino acid, a translation start signal or a translation termination signal. For example, the nucleotide triplets TTA, TTG,

CTT, CTC, CTA and CTG encode the amino acid leucine ("Leu"), TAG, TAA and TGA are translation stop signals and ATG is a translation start signal.

Reading Frame - The grouping of codons during translation of mRNA into amino acid sequences.

During translation, the proper reading frame must be maintained. For example, the sequence 5 GCTGGTTGTAAG may be translated in three reading frames or phases, each of which affords a different amino acid sequence

10 GCT GGT TGT AAG - Ala-Gly-Cys-Lys

G CTG GTT GTA AG - Leu-Val-Val

GC TGG TTG TAA G - Trp-Leu- (STOP) .

Polypeptide - A linear array of amino acids connected one to the other by peptide bonds between the

15 alpha-amino and carboxy groups of adjacent amino acids.

Genome - The entire DNA of a cell or a virus. It includes *inter alia* the structural genes coding for the polypeptides of the cell or virus, as well as its operator, promoter and ribosome binding and interaction sequences, including sequences such as the Shine-Dalgarno sequences.

Structural Gene - A DNA sequence which encodes through its template or messenger RNA ("mRNA") a

20 sequence of amino acids characteristic of a specific polypeptide.

Transcription - The process of producing mRNA from a structural gene.

Translation - The process of producing a polypeptide from mRNA.

Expression - The process undergone by a structural gene to produce a polypeptide. It is a combination of transcription and translation.

25 Plasmid - A non-chromosomal double-stranded DNA sequence comprising an intact "replicon" such that the plasmid is replicated in a host cell. When the plasmid is placed within a unicellular organism, the characteristics of that organism may be changed or transformed as a result of the DNA of the plasmid. For example, a plasmid carrying the gene for tetracycline resistance (Tet^R) transforms a cell previously sensitive to tetracycline into one which is resistant to it. A cell transformed by a plasmid is called a "transformant".

30 Phage or Bacteriophage - Bacterial virus, many of which consist of DNA sequences encapsulated in a protein envelope or coat ("capsid protein").

35 Cloning Vehicle - A plasmid, phage DNA or other DNA sequence which is capable of replicating in a host cell, which is characterized by one or a small number of endonuclease recognition sites at which such DNA sequences may be cut in a determinable fashion without attendant loss of an essential biological function of the DNA, e.g., replication, production of coat proteins or loss of promoter or binding sites, and which contains a marker suitable for use in the identification of transformed cells, e.g., tetracycline resistance or ampicillin resistance. A cloning vehicle is often called a vector.

40 Cloning - The process of obtaining a population of organisms or DNA sequences derived from one such organism or sequence by asexual reproduction.

45 Recombinant DNA Molecule or Hybrid DNA - A molecule consisting of segments of DNA from different genomes which have been joined end-to-end outside of living cells and have the capacity to infect some host cell and be maintained therein.

cDNA Expression Vector - A procaroytic cloning vehicle which also contains sequences of nucleotides that facilitate expression of cDNA sequences in eucaroytic cells. These nucleotides include sequences that 45 function as eucaryotic promoter, alternative splice sites and polyadenylation signals.

Transformation/Transfection - DNA or RNA is introduced into cells in such a way as to allow gene expression. "Infected" referred to herein concerns the introduction of RNA or DNA by a viral vector into the host.

"Injected" referred to herein concerns the microinjection (use of a small syringe) of DNA into a cell.

50 CEA antigen family (CEA gene family) - a set of genes (gene family) and their products (antigen family) that share nucleotide sequences homologous to partial cDNA LV-7 (CEA-(a)) and as a result of these similarities also share a subset of their antigenic epitopes. Examples of the CEA antigen family include CEA (=CEA-(b)), transmembrane CEA (TMCEA) = CEA-(c) and normal crossreacting antigen NCA (=CEA-(d)).

SUMMARY OF THE INVENTION

The present invention concerns the following DNA sequences designated as TM-2 (CEA-(e)), TM-3 (CEA-(f)), TM-4 (CEA-(g)), KGCEA1 and KGCEA2, which code for CEA antigen family peptide sequences:

5

SEQUENCE AND TRANSLATION OF cDNA OF TM-2

10

10	30	50
----	----	----

CAGCCGTGCTCGAACGCGTCCCTGGAGCCCAAGCTCTCCTCCACAGGTGAAGACAGGGCCA

15

70	90	110
----	----	-----

GCAGGAGACACCATGGGGCACCTCTCAGCCCCACTTCACAGAGTGCCTGTACCCCTGGCAG
MetGlyHisLeuSerAlaProLeuHisArgValArgValProTrpGln

20

130	150	170
-----	-----	-----

GGGCTTCTGCTCACAGCCTCACTTCTAACCTCTGGAACCCGCCACCACTGCCAGCTC
GlyLeuLeuLeuThrAlaSerLeuLeuThrPheTrpAsnProProThrThrAlaGlnLeu

25

190	210	230
-----	-----	-----

ACTACTGAATCCATGCCATTCAATGTTGCAGAGGGGAAGGAGGTTCTTCTCCTGTCCAC
ThrThrGluSerMetProPheAsnValAlaGluGlyLysGluValLeuLeuLeuValHis

30

250	270	290
-----	-----	-----

AATCTGCCCAAGCAACTTTTGCTACAGCTGGTACAAAGGGAAAGAGTGGATGGCAAC
AsnLeuProGlnGlnLeuPheGlyTyrSerTrpTyrLysGlyGluValAspGlyAsn

35

310	330	350
-----	-----	-----

CGTCAAATTGAGGATATGCAATAGGAACCTCAACAAGCTACCCAGGGCCCGCAAACAGC
ArgGlnIleValGlyTyrAlaIleGlyThrGlnGlnAlaThrProGlyProAlaAsnSer

40

370	390	410
-----	-----	-----

GGTCGAGAGACAATATACCCCAATGCATCCCTGCTGATCCAGAACGTACCCAGAAATGAC
GlyArgGluThrIleTyrProAsnAlaSerLeuLeuIleGlnAsnValThrGlnAsnAsp

45

430	450	470
-----	-----	-----

ACAGGATTCTATACCCCTACAAAGTCATAAAAGTCAGATCTTG'TGAATGAAGAAGCAACTGGCA
ThrGlyPheTyrThrLeuGlnValIleLysSerAspLeuValAsnGluGluAlaThrGly

55

490 510 530
 CAGTTCCATGTATAACCGGAGCTGCCAAGCCCTCCATCTCCAGCAACAACCTCCAACCT
 5 GlnPheHisValTyrProGluLeuProLysProSerIleSerSerAsnAsnSerAsnPro

 550 570 590
 GTGGAGGACAAGGATGCTGTGCCCTCACCTGTGAACCTGAGACTCAGGACACAACCTAC
 10 ValGluAspLysAspAlaValAlaPheThrCysGluProGluThrGlnAspThrThrTyr

 610 630 650
 CTGTGGTGGATAAACAAATCAGAGCCTCCCGTCAGTCCCAGGCTGCAGCTGTCCAATGGC
 15 LeuTrpTrpIleAsnAsnGlnSerLeuProValSerProArgLeuGlnLeuSerAsnGly

 670 690 710
 AACAGGACCCCTCACTCTACTCAGTGTACAAAGGAATGACACAGGACCCATGAGTGTGAA
 20 AsnArgThrLeuThrLeuLeuSerValThrArgAsnAspThrGlyProTyrGluCysGlu

 730 750 770
 ATACAGAACCCAGTGAGTGCAGACCCGAGTGACCCAGTCACCTTGAATGTCACCTATGCC
 25 IleGlnAsnProValSerAlaAsnArgSerAspProValThrLeuAsnValThrTyrGly

 790 810 830
 CCGGACACCCCCACCATTCCCTTCAGACACCTATTACCGTCCAGGGGAAACCTCAGC
 30 ProAspThrProThrIleSerProSerAspThrTyrTyrArgProGlyAlaAsnLeuSer

 850 870 890
 CTCTCCTGCTATGCAGCCTCTAACCCACCTGCACAGTACTCCTGGCTTATCAATGCAACA
 35 LeuSerCysTyrAlaAlaSerAsnProProAlaGlnTyrSerTrpLeuIleAsnGlyThr

 910 930 950
 TTCCAGCAAAGCACACAAGAGCTTTATCCCTAACATCACTGTGAATAATAGTGGATCC
 40 PheGlnGlnSerThrGlnGluLeuPheIleProAsnIleThrValAsnAsnSerGlySer

 970 990 1010
 TATACCTGGCACGCCAATAACTCAGTCACTGGCTGCAACAGGACACAGTCAAGACGATC
 45 TyrThrCysHisAlaAsnAsnSerValThrGlyCysAsnArgThrThrValLysThrIle

	1030	1050	1070
5	ATAGTCACTGATAATGCTCTACCACAAGAAAATGGCCTCTCACCTGGGCCATTGCTGGC IleValThrAspAsnAlaLeuProGlnGluAsnGlyLeuSerProGlyAlaIleAlaGly		
	1090	1110	1130
10	ATTGTGATTGGAGTACTGGCCCTGGTTGCTCTGATAGCAGTAGCCCTGGCATGTTTCTG IleValIleGlyValValAlaLeuValAlaLeuIleAlaValAlaLeuAlaCysPheLeu		
	1150	1170	1190
15	CATTTGGAAAGACCGGCAGGGCAAGCGACCAGCGTGTATCTCACAGAGCACAAACCTCA HisPheGlyLysThrGlyArgAlaSerAspGlnArgAspLeuThrGluHisLysProSer		
	1210	1230	1250
20	GTCTCCAACCACACTCAGGACCACCTCCAATGACCCACCTAACAAAGATGAATGAAGTTACT ValSerAsnHisThrGlnAspHisSerAsnAspProProAsnLysMetAsnGluValThr		
	1270	1290	1310
25	TATTCTACCTGAACCTTGAAAGCCCAGCAACCCACACAACCAACTTCAGCCTCCCCATCC TyrSerThrLeuAsnPheGluAlaGlnGlnProThrGlnProThrSerAlaSerProSer		
	1330	1350	1370
30	CTAACAGCCACAGAAATAATTATTCAGAACTAAAAAGCAGTAATGAAACCTGTCCCTGC LeuThrAlaThrGluIleIleTyrSerGluValLysLysGln		
	1390	1410	1430
35	TCACTGCAGTGCTGATGTATTCAGTCACCTCTCACCTCATCACTAGGAGATTCCCTTCCC		
	1450	1470	1490
40	CTGTAGGGTAGAGGGGTGGGACAGAAACAACCTTCTCCTACTCTTCTTCTTAATAGGC		
	1510	1530	1550
45	ATCTCCAGGCTGCCCTGGTCACTGCCCTCTCAGTGTCATAGATGAAAGTACATTGGG		
	1570	1590	1610
50	AGTCTGTAGGAAACCAACCTCTTGTCATGAAAATTGGCAAGCTGACTTTGGGAAAG		

1630 1650 1670
 AGGGACCAGAACTTCCCTCCCTTCCCTTTCCAAACCTGGACTTGTAAAAACTTGCC
 5

1690 1710 1730
 TGTCAGAGCACTCATTCCCTTCCACCCCCAGTCCTGTCTATCACTCTAATTGGATTT
 10

1750 1770 1790
 GCCATAGCCTTGAGGTTATGCTCTTCCATTAAGTACATGTGCCAGGAAACACCGAGAG
 15

1810 1830 1850
 AGAGAAAGTAAACGGCAGTAATGCTCTCCTATTCCTCAAAGCCTTGTGAACTAGCA
 20

1870 1890 1910
 AAGAGAAGAAAATCAAATATATAACCAATAGTGAATGCCACAGGTTGTCCACTGTCA
 25

1930 1950 1970
 GGTTGTCTACCTGTAGGATCAGGGTCTAACGCACCTGGTGCTTAGCTAGAATACCACCTA
 30

1990 2010 2030
 ATCCTTCTGGCAAGCCTGCTTCAGAGAAACCACTAGAAGCAACTAGGAAAATCACTTG
 35

2050 2070 2090
 CCAAAATCCAAGGCAATTCTGATGGAAAATGCAAAAGCACATATATGTTTAATATCTT
 40

2110 2130 2150
 TATGGGCTCTGTTCAAGGCAGTGCTGAGAGGGAGGGTTATAGCTTCAGGAGGGAACCAG
 45

2170 2190 2210
 CTTCTGATAAACACAAATCTGCTAGGAACCTGGAAAGGAATCAGAGAGCTGCCCTTCAGC

2230	2250	2270	
GATTATTTAAATTGTTAAAGAATACACAATTGGGGTATTGGGATTTTCTCCTTTCTC			
5	2290	2310	2330
TGAGACATTCCACCATTAAATTGGTAACTGCTTATTTATGTGAAAAGGGTTATT			
10	2350	2370	2390
ACTTAGCTTAGCTATGTCAGCCAATCCGATTGCCTTAGGTGAAAAGAACCACCGAAATCC			
15	2410	2430	2450
CTCAGGTCCCTGGTCAGGAGCCTCTCAAGATTTTTGTCAGAGGCTCCAAATAGAAA			
20	2470	2490	2510
ATAAGAAAAGGTTTCTTCATTCAATGGCTAGAGCTAGATTAACTCAGTTCTAGGCACC			
25	2530	2550	2570
TCAGACCAATCATCAACTACCATTCTATTCCATGTTGCACCTGTGCATTCTGTTGC			
30	2590	2610	2630
CCCCATTCACTTGTCAAGGAAACCTTGCCCTCTGCTAAGGTGTATTGGCTTGAGAAG			
35	2650	2670	2690
TGGGAGCACCCCTACAGGGACACTATCACTCATGCTGGTGGCATTTAGCTAGAAAG			
40	2710	2730	2750
CTGCACTGGTCTAATGCCCTTGGAAATGGGGCTGTGAGGAGGAGGATTATAACCTAG			
45	2770	2790	2810
GCCTAGCCTTTAACAGCCTCTGAAATTATCTTTCTATGGGTCTATAAAATGT			
50	2830	2850	2870
ATCTTATAATAACGGAAACGACAGGAGGAAAGACAGGCAAAATGTAACCTCTCACCCAGTC			

2890 2910 2930
TCTACACAGATGGAATCTCTTGGGCTAAGAGAAAGGTTTATTCTATATTGCTTACCT

5

2950 2970 2990
GATCTCATGTTAGGCCTAACAGAGGCTTCTCCAGGAGGATTAGCTTGGAGTTCTCTATACT

10

3010 3030 3050
CAGGTACCTCTTCAGGGTTTCTAACCCCTGACACGGACTGTGCATACTTTCCCTCATCC

15

3070 3090 3110
ATGCTGTGCTGTGTTATTTAATTTTCCTGGCTAACGATCATGTCATGAAATTATGATGAAA

20

3130 3150 3170
ATTATTCTATGTTTTATAATAAAAAATAATATCAGACATCGAAAAAA

25

30

35

40

45

50

55

SEQUENCE AND TRANSLATION OF cDNA OF TM-3

5

430 450 470
 5 ACAGGATTCTACACCCTACAAGTCATAAAAGTCAGATCTTGTGAATGAAGAAGCAACTGGA
 ThrGlyPheTyrThrLeuGlnValIleLysSerAspLeuValAsnGluGluAlaThrGly

 490 510 530
 10 CAGTTCCATGTATAACCCGGAGCTGCCAAGCCCTCATCTCCAGCAACAACTCCAACCC
 GlnPheHisValTyrProGluLeuProLysProSerIleSerSerAsnAsnSerAsnPro

 550 570 590
 15 GTGGAGGACAAGGATGCTGTGGCCTTCACCTGTGAACCTGAGACTCAGGACACAAACCTAC
 ValGluAspLysAspAlaValAlaPheThrCysGluProGluThrGlnAspThrThrTyr

 610 630 650
 20 CTGTGGTGGATAAACAAATCAGAGCCTCCGGTCAGTCCCAGGCTGCCAGCTGTCCAATGGC
 LeuTrpTrpIleAsnAsnGlnSerLeuProValSerProArgLeuGlnLeuSerAsnGly

 670 690 710
 25 AACAGGACCCCTCACTCTACTCAGTGTACAAGGAATGACACAGGACCCATGAGTGTGAA
 AsnArgThrLeuThrLeuSerValThrArgAsnAspThrGlyProTyrGluCysGlu

 730 750 770
 30 ATACAGAACCCAGTGAGTGCAGACCGCAGTGACCCAGTCACCTTGAATGTCACCTATGGC
 IleGlnAsnProValSerAlaAsnArgSerAspProValThrLeuAsnValThrTyrGly

 790 810 830
 35 CCGGACACCCCCCACCATTCCCTTCAGACACCTATTACCGTCCAGGGGCAAACCTCAGC
 ProAspThrProThrIleSerProSerAspThrTyrArgProGlyAlaAsnLeuSer

 40

 45

 50

 55

850

870

890

5 CTCTCCTGCTATGCAGCCTCTAACCCACCTGCACAGTACTCCTGGCTTATCAATGGAACA
 LeuSerCysTyrAlaAlaSerAsnProProAlaGlnTyrSerTrpLeuIleAsnGlyThr

910

930

950

10 TTCCAGCAAAGCACACAAGAGCTCTTATCCCTAACATCACTGTGAATAATAGTGGATCC
 PheGlnGlnSerThrGlnGluLeuPheIleProAsnIleThrValAsnAsnSerGlySer

970

990

1010

15 TATACCTGCCACGCCAATAACTCAGTCACTGGCTGCAACAGGACCACAGTCAGACGATC
 TyrThrCysHisAlaAsnAsnSerValThrGlyCysAsnArgThrThrValLysThrIle

1030

1050

1070

20 ATAGTCACTGAGCTAACGTCAGTAGTAGCAAAGCCCCAAATCAAAGCCAGCAAGACCACA
 IleValThrGluLeuSerProValValAlaLysProGlnIleLysAlaSerLysThrThr

1090

1110

1130

25 GTCACAGGAGATAAGGACTCTGTGAACCTGACCTGCTCCACAAATGACACTGGAATCTCC
 ValThrGlyAspLysAspSerValAsnLeuThrCysSerThrAsnAspThrGlyIleSer

1150

1170

1190

30 ATCCGTTGGTTCTCAAAAACCAGAGTCTCCGTCTCGGAGAGGGATGAAGCTGTCCAG
 IleArgTrpPhePheLysAsnGlnSerLeuProSerSerGluArgMetLysLeuSerGln

1210

1230

1250

35 GGCAACACCACCCCTCAGCATAAACCCCTGTCAAGAGGGAGGATGCTGGGACGTATTGGTGT
 GlyAsnThrThrLeuSerIleAsnProValLysArgGluAspAlaGlyThrTyrTrpCys

40

45

50

55

1270 1290 1310
 5 GAGGTCTCAACCCAATCAGTAAGAACCAAGCGACCCATCATGCTGAACGTAAACTAT
 GluValPheAsnProIleSerLysAsnGlnSerAspProIleMetLeuAsnValAsnTyr

 1330 1350 1370
 10 AATGCTCTACCACAAGAAAATGGCCTCTCACCTGGGCCATTGCTGGCATTGTGATTGGA
 AsnAlaLeuProGlnGluAsnGlyLeuSerProGlyAlaIleAlaGlyIleValIleGly

 1390 1410 1430
 15 GTAGTGGCCCTGGTTGCTCTGATAGCAGTAGCCCTGGCATGTTTCTGCATTCGGGAAG
 ValValAlaLeuValAlaLeuIleAlaValAlaLeuAlaCysPheLeuHisPheGlyLys

 1450 1470 1490
 20 ACCGGCAGCTCAGGACCACCCAATGACCCACCTAACAAAGATGAATGAAGTTACTTATT
 ThrGlySerSerGlyProLeuGln

 1510 1530 1550
 25 TACCCCTGAACTTGAAGGCCAGCAACCCACACAACCAACTTCAGCCTCCCCATCCCTAAC

 1570 1590 1610
 30 AGCCACAGAAATAATTATTCAGAAGTAAAAAAGCAGTAATGAAACCTGAAAAAAAAAAA

 1630
 35 AAAAAAAA

 40

 45

 50

 55

SEQUENCE AND TRANSLATION OF cDNA OF TM-4

5	10	30	50
	CAGCCGTGCTCGAAGCGTTCTGGAGCCCAAGCTCTCCTCCACAGGTGAAGACAGGGCCA		
10	70	90	110
	GCAGGAGACACCATGGGCACCTCTCAGCCCCACTTCACAGAGTGCAGTGTACCCCTGGCAG MetGlyHisLeuSerAlaProLeuHisArgValArgValProTrpGln		
15	130	150	170
	GGGCTTCTGCTCACAGCCTCACTTCTAACCTTCTCGAACCCGCCACCACTGCCAGCTC GlyLeuLeuLeuThrAlaSerLeuLeuThrPheTrpAsnProProThrThrAlaGlnLeu		
20	190	210	230
	ACTACTGAATCCATGCCATTCAATGTTGCAGAGGGAAAGGAGGTTCTTCTCCTTGTCAC ThrThrGluSerMetProPheAsnValAlaGluGlyLysGluValLeuLeuLeuValHis		
25	250	270	290
	AATCTGCCCAAGCAACTTTTGCTACAGCTGGTACAAAGGGAAAGACTGGATGGCAAC AsnLeuProGlnGlnLeuPheGlyTyrSerTrpTyrLysGlyGluArgValAspGlyAsn		
30	310	330	350
35	CGTCAAATTGTAGGATATGCAATAGGAACTCAACAAGCTACCCAGGGCCCCGAAACAGC ArgGlnIleValGlyTyrAlaIleGlyThrGlnGlnAlaThrProGlyProAlaAsnSer		
40	370	390	410
	GGTCGAGAGACAATATACCCCAATGCATCCCTGCTGATCCAGAACGTCACCCAGAAATGAC GlyArgGluThrIleTyrProAsnAlaSerLeuLeuIleGlnAsnValThrGlnAsnAsp		
45	430	450	470
	ACAGGATTCTACACCCCTACAAAGTCATAAAAGTCAGATCTTGTGAATGAAAGCAACTCGA ThrGlyPheTyrThrLeuGlnValIleLysSerAspLeuValAsnGluGluAlaThrGly		

490	510	530
CAGTTCCATGTATAACCGGAGCTGCCAAGCCCTCCATCTCCAGCAACA 5 GlnPheHisValTyrProGluLeuProLysProSerIleSerSerAsnAsnSerAsnPro		
550	570	590
GTGGAGGACAAGGATGCTGTGGCTTCACCTGTGAACCTGAGACTCAGGACACA 10 ValGluAspLysAspAlaValAlaPheThrCysGluProGluThrGlnAspThrThrTyr		
610	630	650
CTGTGGTGGATAAACAAATCAGAGCCTCCGGTCAGTCCCAGGCTGCAGCTGTCCAATGGC 15 LeuTrpTrpIleAsnAsnGlnSerLeuProValSerProArgLeuGlnLeuSerAsnGly		
670	690	710
AACAGGACCCCTCACTCTACTCAGTGTACAAGGAATGACACAGGACCCATTGAGTGTGAA 20 AsnArgThrLeuThrLeuLeuSerValThrArgAsnAspThrGlyProTyrGluCysGlu		
730	750	770
ATACAGAACCCAGTGAGTGCAGACCCAGTGACCCAGTCACCTTGAATGTCACCTATGGC 25 IleGlnAsnProValSerAlaAsnArgSerAspProValThrLeuAsnValThrTyrGly		
790	810	830
CCGGACACCCCCACCATTTCCCTTCAGACACCTATTACCGTCCAGGGGCAACCTCAGC 30 ProAspThrProThrIleSerProSerAspThrTyrTyrArgProGlyAlaAsnLeuSer		
850	870	890
CTCTCCTGCTATGCAGCCTCTAACCCACCTGCACAGTACTCCTGGCTTATCAATGGAA 35 CA LeuSerCysTyrAlaAlaSerAsnProProAlaGlnTyrSerTrpLeuIleAsnGlyThr		
910	930	950
TTCCAGCAAAGCACACAAGAGCTTTATCCCTAACATCACTGTGAATAATAGTGGATCC 40 PheGlnGlnSerThrGlnGluLeuPheIleProAsnIleThrValAsnAsnSerGlySer		
970	990	1010
TATACCTGCCACGCCAATAACTCAGTCACTGGCTGCAACAGGACCAAGTCAGACGATC 45 TyrThrCysHisAlaAsnAsnSerValThrGlyCysAsnArgThrThrValLysThrIle		

50

55

	1030	1050	1070
5	ATAGTCACTGATAATGCTCTACCACAAGAAAATGGCCTCTCACCTGGGCCATTGCTGGC IleValThrAspAsnAlaLeuProGlnGluAsnGlyLeuSerProGlyAlaIleAlaGly		
	1090	1110	1130
10	ATTGTGATTGGAGTAGTGGCCCTGGTTGCTCTGATAGCAGTAGCCCTGGCATGTTTCTG IleValIleGlyValValAlaLeuValAlaLeuIleAlaValAlaLeuAlaCysPheLeu		
	1150	1170	1190
15	CATTCGGGAAGACCGGCAGCTCAGGACCCTCAATGACCCACCTAACAAAGATGAATGA HisPheGlyLysThrGlySerSerGlyProLeuGln		
	1210	1230	1250
20	AGTTACTTATTCTACCCCTGAACCTTGAGGCCAGCAACCCACACAAACCAACTTCAGCCTC		
	1270	1290	1310
25	CCCATCCCTAACAGCCACAGAAATAATTATTAGAAGTAAAAAGCAGTAAATGAAACCT		
	1330		
30	GAAAAAAAAAAAAAA		

The present invention is also directed to a replicable recombinant cloning vehicle ("vector") having an insert comprising a nucleic acid, e.g., DNA, which comprises a base sequence which codes for a CEA peptide or a base sequence hybridizable therewith.

This invention also relates to a cell that is transformed/transfected, infected or injected with the above described replicable recombinant cloning vehicle or nucleic acid hybridizable with the aforementioned cDNA. Thus the invention also concerns the transfection of cells using free nucleic acid, without the use of a cloning vehicle.

Still further, the present invention concerns a polypeptide expressed by the above described transfected, infected or injected cell, which polypeptide exhibits immunological cross-reactivity with a CEA, as well as labelled forms of the polypeptide. The invention also relates to polypeptides having an amino acid sequence, i.e., synthetic peptides, or the expression product of a cell that is transfected, injected, infected with the above described replicable recombinant cloning vehicles, as well as labelled forms thereof. Stated otherwise, the present invention concerns a synthetic peptide having an amino acid sequence corresponding to the entire amino acid sequence or a portion thereof having no less than five amino acids of the aforesaid expression product.

The invention further relates to an antibody preparation specific for the above described polypeptide.

Another aspect of the invention concerns an immunoassay method for detecting CEA or a functional equivalent thereof in a test sample comprising

- (a) contacting the sample with the above described antibody preparation, and
- (b) determining binding thereof to CEA in the sample.

The invention also is directed to a nucleic acid hybridization method for detecting a CEA or a related nucleic acid (DNA or RNA) sample in a test sample comprising

(a) contacting the test sample with a nucleic acid probe comprising a nucleic acid, which comprises a base sequence which codes for a CEA peptide sequence or a base sequence that is hybridizable therewith, and

- (b) determining the formation of the resultant hybridized probe.

The present invention also concerns a method for detecting the presence of carcinoembryonic antigen or a functional equivalent thereof in an animal or human patient *in vivo* comprising

- a) introducing into said patient a labeled (e.g., a radio-opaque material that can be detected by X-rays, radiolabeled or labeled with paramagnetic materials that can be detected by NMR) antibody preparation according to the present invention and
- 5 b) detecting the presence of such antibody preparation in the patient by detecting the label.

In another aspect, the present invention relates to the use of an antibody preparation according to the present invention for therapeutic purposes, namely, attaching to an antibody preparation radionuclides, toxins or other biological effectors to form a complex and introducing an effective amount of such complex 10 into an animal or human patient, e.g., by injection or orally. The antibody complex would attach to CEA in a patient and the radionuclide, toxin or other biological effector would serve to destroy the CEA expressing cell.

BRIEF DESCRIPTION OF THE DRAWINGS

15 Fig. 1 is a schematic representation of the transmembrane CEA's

DETAILED DESCRIPTION OF THE INVENTION

20 In the parent application 87111/68, published as EP-A-263 933, applicants described the following CEA's:

	ATCC No.
25	CEA-(a) partial CEA (pcLV7)
	CEA-(b) full coding CEA (pc 15LV7)
	CEA-(c) TM-1 (FL-CEA; pc 19-22)
	CEA-(d) NCA (pcBT 20)
	67709
	67710
	67711

30 In the present application, applicants described the following CEA's:

	ATCC No.
35	CEA-(e) TM-2 (pc E22)
	CEA-(f) TM-3 (pc HT-6)
	CEA-(g) TM-4.
	67712
	67708

ATCC Nos. 67708, 67709, 67710, 67711 and 67712 were all deposited with the American Type Culture

40 Collection on May 25, 1988.

45

50

55

The sequences for CEA-(a), CEA-(b), CEA-(c) and CEA-(d) are given hereinbelow:

CEA-(a):

5

10

15

20

25

30

35

40

(b)

45

50

55

10 20 30 40 50

C ACC ATG GAG TCT CCC TCG GCC CCT CTC CAC AGA TGG TGC ATC CCC TGG CAG AGG CTC
 Met Glu Ser Pro Ser Ala Pro Leu His Arg Trp Cys Ile Pro Trp Gln Arg Leu

	60	70	80	90	100	110														
5	*	*	*	*	*	*														
	CTG	CTC	ACA	GCC	TCA	CTT	CTA	ACC	TTC	TGG	AAC	CCG	CCC	ACC	ACT	GCC	AAG	CTC	ACT	
	Leu	Leu	Thr	Ala	Ser	Leu	Leu	Thr	Phe	Trp	Asn	Pro	Pro	Thr	Thr	Ala	Lys	Leu	Thr	
																				1 2 3
10	120	130	140	150	160	170														
	*	*	*	*	*	*														
	ATT	GAA	TCC	ACG	CCG	TTC	AAT	GTC	GCA	GAG	GGG	AAG	GAG	GTG	CTT	CTA	CTT	GTC	CAC	
	Ile	Glu	Ser	Thr	Pro	Phe	Asn	Val	Ala	Glu	Gly	Lys	Glu	Val	Leu	Leu	Leu	Val	His	
	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	
15	180	190	200	210	220															
	*	*	*	*	*															
	AAT	CTG	CCC	CAG	CAT	CTT	TTT	GGC	TAC	AGC	TGG	TAC	AAA	GGT	GAA	AGA	GTG	GAT	GGC	
	Asn	Leu	Pro	Gln	His	Leu	Phe	Gly	Tyr	Ser	Trp	Tyr	Lys	Gly	Glu	Arg	Val	Asp	Gly	
	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	
20	230	240	250	260	270	280														
	*	*	*	*	*	*														
	AAC	CGT	CAA	ATT	ATA	GGA	TAT	GTA	ATA	GGA	ACT	CAA	CAA	GCT	ACC	CCA	GGG	CCC	GCA	
25	Asn	Arg	Gln	Ile	Ile	Gly	Tyr	Val	Ile	Gly	Thr	Gln	Gln	Ala	Thr	Pro	Gly	Pro	Ala	
	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	
30	290	300	310	320	330	340														
	*	*	*	*	*	*														
	TAC	AGT	GGT	CGA	GAG	ATA	ATA	TAC	CCC	AAT	GCA	TCC	CTG	CTG	ATC	CAG	AAC	ATC	ATC	
	Tyr	Ser	Gly	Arg	Glu	Ile	Ile	Tyr	Pro	Asn	Ala	Ser	Leu	Leu	Ile	Gln	Asn	Ile	Ile	
	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	
35	350	360	370	380	390	400														
	*	*	*	*	*	*														
	CAG	AAT	GAC	ACA	GGA	TTC	TAC	ACC	CTA	CAC	GTC	ATA	AAG	TCA	GAT	CTT	GTG	AAT	GAA	
	Gln	Asn	Asp	Thr	Gly	Phe	Tyr	Thr	Leu	His	Val	Ile	Lys	Ser	Asp	Leu	Val	Asn	Glu	
	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	
40	410	420	430	440	450															
	*	*	*	*	*															
	GAA	GCA	ACT	GGC	CAG	TTC	CGG	GTA	TAC	CCG	GAG	CTG	CCC	AAG	CCC	TCC	ATC	TCC	AGC	
	Glu	Ala	Thr	Gly	Gln	Phe	Arg	Val	Tyr	Pro	Glu	Leu	Pro	Lys	Pro	Ser	Ile	Ser	Ser	
	99	101	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	
45	460	470	480	490	500	510														
	*	*	*	*	*	*														
	AAC	AAC	TCC	AAA	CCC	GTG	GAG	GAC	AAG	GAT	GCT	GTG	GCC	TTC	ACC	TGT	GAA	CCT	GAG	
50	Asn	Asn	Ser	Lys	Pro	Val	Glu	Asp	Lys	Asp	Ala	Val	Ala	Phe	Thr	Cys	Glu	Pro	Glu	
	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	

EP 0 346 710 B1

	520	530	540	550	560	570
5	ACT CAG GAC GCA ACC TAC CTG TGG TGG GTA AAC AAT CAG AGC CTC CCG GTC AGT CCC Thr Gln Asp Ala Thr Tyr Leu Trp Trp Val Asn Asn Gln Ser Leu Pro Val Ser Pro 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155					
10	580	590	600	610	620	
	AGG CTG CAG CTG TCC AAT GGC AAC AGG ACC CTC ACT CTA TTC AAT GTC ACA AGA AAT Arg Leu Gln Leu Ser Asn Gly Asn Arg Thr Leu Thr Leu Phe Asn Val Thr Arg Asn 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174					
15	630	640	650	660	670	680
	GAA CAA GCA AGC TAC AAA TGT GAA ACC CAG AAC CCA GTG AGT GCC AGG CGC AGT GAT Glu Gln Ala Ser Tyr Lys Cys Glu Thr Gln Asn Pro Val Ser Ala Arg Arg Ser Asp 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193					
20	690	700	710	720	730	740
25	TCA GTC ATC CTG AAT GTC CTC TAT GGC CCG GAT GCC CCC ACC ATT TCC CCT CTA AAC Ser Val Ile Leu Asn Val Leu Tyr Gly Pro Asp Ala Pro Thr Ile Ser Pro Leu Asn 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212					
30	750	760	770	780	790	
	ACA TCT TAC AGA TCA GGG GAA AAT CTG AAC CTC TCC TGC CAC GCA GCC TCT AAC CCA Thr Ser Tyr Arg Ser Gly Glu Asn Leu Asn Leu Ser Cys His Ala Ala Ser Asn Pro 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231					
35	800	810	820	830	840	850
	CCT GCA CAG TAC TCT TGG TTT GTC AAT GGG ACT TTC CAG CAA TCC ACC CAA GAG CTC Pro Ala Gln Tyr Ser Trp Phe Val Asn Gly Thr Phe Gln Gln Ser Thr Gln Glu Leu 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250					
40	860	870	880	890	900	910
45	TTT ATC CCC AAC ATC ACT GTG AAT AAT AGT GGA TCC TAT ACG TGC CAA GCC CAT AAC Phe Ile Pro Asn Ile Thr Val Asn Asn Ser Gly Ser Tyr Thr Cys Gln Ala His Asn 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269					
50	920	930	940	950	960	970
	TCA GAC ACT GGC CTC AAT AGG ACC ACA GTC ACG ACG ATC ACA GTC TAT GCA GAG CCA Ser Asp Thr Gly Leu Asn Arg Thr Thr Val Thr Ile Thr Val Tyr Ala Glu Pro 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288					

EP 0 346 710 B1

	980	990	1000	1010	1020
5	CCC AAA CCC TTC ATC ACC AGC AAC AAC TCC AAC CCC GTG GAG GAT GAG GAT GCT GTA Pro Lys Pro Phe Ile Thr Ser Asn Asn Ser Asn Pro Val Glu Asp Glu Asp Ala Val 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307				
10	1030	1040	1050	1060	1070
	GCC TTA ACC TGT GAA CCT GAG ATT CAG AAC ACA ACC TAC CTG TGG TGG GTA AAT AAT Ala Leu Thr Cys Glu Pro Glu Ile Gln Asn Thr Thr Tyr Leu Trp Trp Val Asn Asn 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326				
15	1090	1100	1110	1120	1130
	CAG AGC CTC CCG GTC AGT CCC AGG CTG CAG CTG TCC AAT GAC AAC AGG ACC CTC ACT Gln Ser Leu Pro Val Ser Pro Arg Leu Gln Leu Ser Asn Asp Asn Arg Thr Leu Thr 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345				
20	1150	1160	1170	1180	1190
	CTA CTC AGT GTC ACA AGG AAT GAT GTA GGA CCC TAT GAG TGT GGA ATC CAG AAC GAA Leu Leu Ser Val Thr Arg Asn Asp Val Gly Pro Tyr Glu Cys Gly Ile Gln Asn Glu 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364				
25	1200	1210	1220	1230	1240
	TTA AGT GTT GAC CAC AGC GAC CCA GTC ATC CTG AAT GTC CTC TAT GGC CCA GAC GAC Leu Ser Val Asp His Ser Asp Pro Val Ile Leu Asn Val Leu Tyr Gly Pro Asp Asp 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383				
30	1260	1270	1280	1290	1300
	CCC ACC ATT TCC CCC TCA TAC ACC TAT TAC CGT CCA GGG GTG AAC CTC AGC CTC TCC Pro Thr Ile Ser Pro Ser Tyr Thr Tyr Arg Pro Gly Val Asn Leu Ser Leu Ser 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402				
35	1320	1330	1340	1350	1360
	TGC CAT GCA GCC TCT AAC CCA CCT GCA CAG TAT TCT TGG CTG ATT GAT GGG AAC ATC Cys His Ala Ala Ser Asn Pro Pro Ala Gln Tyr Ser Trp Leu Ile Asp Gly Asn Ile 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421				
40	1370	1380	1390	1400	1410
	CAG CAA CAC ACA CAA GAG CTC TTT ATC TCC AAC ATC ACT GAG AAG AAC AGC GGA CTC Gln Gln His Thr Gln Glu Leu Phe Ile Ser Asn Ile Thr Glu Lys Asn Ser Gly Leu 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440				

EP 0 346 710 B1

	1430	1440	1450	1460	1470	1480														
5	TAT	ACC	TGC	CAG	GCC	AAT	AAC	TCA	GCC	AGT	GGC	CAC	AGC	AGG	ACT	ACA	GTC	AAG	ACA	
	Tyr	Thr	Cys	Gln	Ala	Asn	Asn	Ser	Ala	Ser	Gly	His	Ser	Arg	Thr	Thr	Val	Lys	Thr	
	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	
10	1490	1500	1510	1520	1530	1540														
	ATC	ACA	GTC	TCT	GCG	GAC	GTG	CCC	AAG	CCC	TCC	ATC	TCC	AGC	AAC	AAC	TCC	AAA	CCC	
	Ile	Thr	Val	Ser	Ala	Asp	Val	Pro	Lys	Pro	Ser	Ile	Ser	Ser	Asn	Asn	Ser	Lys	Pro	
	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	
15	1550	1560	1570	1580	1590															
	GTG	GAG	GAC	AAG	GAT	GCT	GTG	GCC	TTC	ACC	TGT	GAA	CCT	GAG	GCT	CAG	AAC	ACA	ACC	
	Val	Glu	Asp	Lys	Asp	Asp	Ala	Val	Ala	Phe	Thr	Cys	Glu	Pro	Glu	Ala	Gln	Asn	Thr	Thr
	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	
20	1600	1610	1620	1630	1640	1650														
	TAC	CTG	TGG	TGG	GTA	AAT	GGT	CAG	AGC	CTC	CCA	GTC	AGT	CCC	AGG	CTG	CAG	CTG	TCC	
	Tyr	Leu	Trp	Trp	Val	Asn	Gly	Gln	Ser	Leu	Pro	Val	Ser	Pro	Arg	Leu	Gln	Leu	Ser	
25	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	
30	1660	1670	1680	1690	1700	1710														
	AAT	GGC	AAC	AGG	ACC	CTC	ACT	CTA	TTC	AAT	GTC	ACA	AGA	AAT	GAC	GCA	AGA	GCC	TAT	
	Asn	Gly	Asn	Arg	Thr	Leu	Thr	Leu	Phe	Asn	Val	Thr	Arg	Asn	Asp	Ala	Arg	Ala	Tyr	
	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	
35	1720	1730	1740	1750	1760															
	GTA	TGT	GGA	ATC	CAG	AAC	TCA	GTG	AGT	GCA	AAC	CGC	AGT	GAC	CCA	GTC	ACC	CTG	GAT	
	Val	Cys	Gly	Ile	Gln	Asn	Ser	Val	Ser	Ala	Asn	Arg	Ser	Asp	Pro	Val	Thr	Leu	Asp	
	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	
40	1770	1780	1790	1800	1810	1820														
	GTC	CTC	TAT	GGG	CCG	GAC	ACC	CCC	ATC	ATT	TCC	CCC	CCA	GAC	TCG	TCT	TAC	CTT	TCG	
	Val	Leu	Tyr	Gly	Pro	Asp	Thr	Pro	Ile	Ile	Ser	Pro	Pro	Asp	Ser	Ser	Tyr	Leu	Ser	
	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	
45	1830	1840	1850	1860	1870	1880														
	GGA	GCG	AAC	CTC	AAC	CTC	TCC	TGC	CAC	TCG	GCC	TCT	AAC	CCA	TCC	CCG	CAG	TAT	TCT	
	Gly	Ala	Asn	Leu	Asn	Leu	Ser	Cys	His	Ser	Ala	Ser	Asn	Pro	Ser	Pro	Gln	Tyr	Ser	
50	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	

EP 0 346 710 B1

1890 1900 1910 1920 1930

5 TGG CGT ATC AAT GGG ATA CCG CAG CAA CAC ACA CAA GTT CTC TTT ATC GCC AAA ATC
 Trp Arg Ile Asn Gly Ile Pro Gln Gln His Thr Gln Val Leu Phe Ile Ala Lys Ile
 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611

1940 1950 1960 1970 1980 1990

10 ACG CCA AAT AAT AAC GGG ACC TAT GCC TGT TTT GTC TCT AAC TTG GCT ACT GGC CGC
 Thr Pro Asn Asn Asn Gly Thr Tyr Ala Cys Phe Val Ser Asn Leu Ala Thr Gly Arg
 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630

15 2000 2010 2020 2030 2040 2050

AAT AAT TCC ATA GTC AAG AGC ATC ACA GTC TCT GCA TCT GGA ACT TCT CCT GGT CTC
 Asn Asn Ser Ile Val Lys Ser Ile Thr Val Ser Ala Ser Gly Thr Ser Pro Gly Leu
 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649

20 2060 2070 2080 2090 2100 2110

TCA GCT GGG GCC ACT GTC GGC ATC ATG ATT GGA GTG CTG GTT GGG GTT GCT CTG ATA
 Ser Ala Gly Ala Thr Val Gly Ile Met Ile Gly Val Leu Val Gly Val Ala Leu Ile
 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668

25 2120 2130 2140 2150 2160

30 TAG CAG CCC TGG TGT AGT TTC TTC ATT TCA GGA AGA CTG ACA GTT GTT TTG CTT CTT

35 2170 2180 2190 2200 2210 2220
 CCT TAA AGC ATT TGC AAC AGC TAC AGT CTA AAA TTG CTT CTT TAC CAA GGA TAT TTA

40 2230 2240 2250 2260 2270 2280
 CAG AAA ATA CTC TGA CCA GAG ATC GAG ACC ATC CTA GCC AAC ATC GTG AAA CCC CAT

45 2290 2300 2310 2320 2330
 CTC TAC TAA AAA TAC AAA AAT GAG CTG GGC TTG GTG GCG CGC ACC TGT AGT CCC AGT

50 2340 2350 2360 2370 2380 2390
 TAC TCG GGA GGC TGA GGC AGG AGA ATC GCT TGA ACC CGG GAG GTG GAG ATT GCA GTG

EP 0 346 710 B1

2400 2410 2420 2430 2440 2450

AGC CCA GAT CGC ACC ACT GCA CTC CAG TCT GGC AAC AGA GCA AGA CTC CAT CTC AAA

5

2460 2470 2480 2490 2500

AAG AAA AGA AAA GAA GAC TCT GAC CTG TAC TCT TGA ATA CAA GTT TCT GAT ACC ACT

10

2510 2520 2530 2540 2550 2560

GCA CTG TCT GAG AAT TTC CAA AAC TTT AAT GAA CTA ACT GAC AGC TTC ATG AAA CTG

15

2570 2580 2590 2600 2610 2620

TCC ACC AAG ATC AAG CAG AGA AAA TAA TTA ATT TCA TGG GGA CTA AAT GAA CTA ATG

20

2630 2640 2650 2660 2670 2680

AGG ATA ATA TTT TCA TAA TTT TTT ATT TGA AAT TTT GCT GAT TCT TTA AAT GTC TTG

25

2690 2700 2710 2720 2730

TTT CCC AGA TTT CAG GAA ACT TTT TTT CTT TTA AGC TAT CCA CTC TTA CAG CAA TTT

30

2740 2750 2760 2770 2780 2790
GAT AAA ATA TAC TTT TGT GAA CAA AAA TTG AGA CAT TTA CAT TTT ATC CCT ATG TGG

35

2800 2810 2820 2830
TCG CTC CAG ACT TGG GAA ACT ATT CAT GAA TAT TTA TAT TGT ATG

40

45

50

55

CEA- (c):

5

10

30

50

10 CAGCCGTGCTCGAAGCGTTCTGGAGCCCAAGCTCTCCTCCACAGGTGAAGACAGGGCCA

70

70

90

110

15 GCAGGAGACACCATGGGGCACCTCTCAGCCCCACTTCACAGAGTGCCTGTACCCCTGGCAG
MetGlyHisLeuSerAlaProLeuHisArgValArgValProTrpGln

15

130

150

170

20 GGGCTTCTGCTCACAGCCTCACTTCTAACCTTCTGGAACCCGCCACCACTGCCAGCTC
GlyLeuLeuLeuThrAlaSerLeuLeuThrPheTrpAsnProProThrThrAlaGlnLeu

20

190

210

230

25 ACTACTGAATCCATGCCATTCAATGTTGCAGAGGGAAAGGAGGTTCTCTCCTGTCCAC
ThrThrGluSerMetProPheAsnValAlaGluGlyLysGluValLeuLeuValHis

25

250

270

290

30 AATCTGCCCAAGCAACTTTGGCTACAGCTGGTACAAAGGGAAAGAGTGGATGGCAAC
AsnLeuProGlnGlnLeuPheGlyTyrSerTrpTyrLysGlyGluArgValAspGlyAsn

30

310

330

350

35 CGTCAAATTGTTAGGATATGCAATAGGAACCTCAACAAAGCTACCCAGGGCCCGCAAACAGC
ArgGlnIleValGlyTyrAlaIleGlyThrGlnGlnAlaThrProGlyProAlaAsnSer

35

40

370

390

410

40 GGTGGAGAGACAATATAACCCCAATGCATCCCTGCTGATCCAGAACGTCACCCAGAAATGAC
GlyArgGluThrIleTyrProAsnAlaSerLeuLeuIleGlnAsnValThrGlnAsnAsp

40

430

450

470

45 ACAGGATTCTACACCCCTACAAGTCATAAAGTCAGATCTTGTGAATGAAGAAGCAACTGGA
ThrGlyPheTyrThrLeuGlnValIleLysSerAspLeuValAsnGluGluAlaThrGly

50

490

510

530

55

5 CAGTTCCATGTATACTCCGGAGCTGCCAAGCCCTCCATCTCCAGCAACAACCTCAACCT
 GlnPheHisValTyrProGluLeuProLysProSerIleSerSerAsnAsnSerAsnPro

5

550

570

590

10 GTGGAGGACAAGGATGCTGTGGCCTTCACCTGTGAACCTGAGACTCAGGACACAACCTAC
 ValGluAspLysAspAlaValAlaPheThrCysGluProGluThrGlnAspThrThrTyr

10

610

630

650

15

CTGTGGTGGATAAACAAATCAGAGCCTCCGGTCAGTCCCAGGCTGCAGCTGTCCAATGGC
 LeuTrpTrpIleAsnAsnGlnSerLeuProValSerProArgLeuGlnLeuSerAsnGly

20

670

690

710

AACAGGACCCCTCACTCTACTCAGTGTACAAAGGAATGACACAGGGACCCATTGAGTGTGAA
 AsnArgThrLeuThrLeuLeuSerValThrArgAsnAspThrGlyProTyrGluCysGlu

25

730

750

770

ATACAGAACCCAGTGAGTGCAGACCGCAGTGACCCAGTCACCTTGAATGTCACCTATGGC
 IleGlnAsnProValSerAlaAsnArgSerAspProValThrLeuAsnValThrTyrGly

30

790

810

830

CCGGACACCCCCACCATTCCCCCTCAGACACCTATTACCGTCCAGGGCAAACCTCAGC
 ProAspThrProThrIleSerProSerAspThrTyrArgProGlyAlaAsnLeuSer

35

850

870

890

CTCTCCTGCTATGCAGCCTCTAACCCACCTGCACAGTACTCCTGGTTATCAATGGAACA
 LeuSerCysTyrAlaAlaSerAsnProProAlaGlnTyrSerTrpLeuIleAsnGlyThr

40

910

930

950

TTCCAGCAAAGCACACAAGAGCTTTATCCCTAACATCACTGTGAATAATAGTGGATCC
 PheGlnGlnSerThrGlnGluLeuPheIleProAsnIleThrValAsnAsnSerGlySer

45

970

990

1010

TATACCTGCCACGCCAATAACTCAGTCACTGGCTGCAACAGGACCAAGTCAAGACGATC
 TyrThrCysHisAlaAsnAsnSerValThrGlyCysAsnArgThrThrValLysThrIle

50

1030

1050

1070

55

ATAGTCACTGAGCTAAGTCCAGTAGTAGCAAAGCCCCAAATCAAAGCCAGCAAGACCACA
IleValThrGluLeuSerProValValAlaLysProGlnIleLysAlaSerLysThrThr

5 1090 1110 1130
 GTCACAGGAGATAAGGACTCTGTGAACCTGACCTGCTCCACAAATGACACTGGAATCTCC
 Val Thr Gly Asp Lys Asp Ser Val Asn Leu Thr Cys Ser Thr Asn Asp Thr Gly Ile Ser

10
1150 1170 1190
ATCCGTTGGTTCTTCAAAACCAAGAGTCTCCGTCCTCGGAGAGGATGAAGCTGTCCCCAG
IleArgTrpPhePheLysAsnGlnSerLeuProSerSerGluArgMetLysLeuSerGln
15

1210 1230 1250
 GGCAACACCACCCTCAGCATAAACCTGTCAAGAGGGAGGATGCTGGGACGTATTGGTGT
 GlyAsnThrThrLeuSerIleAsnProValLysArgGluAspAlaGlyThrTyrTrpCys
 20

25 1270 1290 1310
 GAGGTCTCAACCCAATCAGTAAGAACCAAAGCGACCCCCATCATGCTGAACGTAAACTAT
 GluValPheAsnProIleSerLysAsnGlnSerAspProIleMetLeuAsnValAsnTyr

1330	1350	1370
30	AATGCTCTACCACAAGAAAATGGCCTCTCACCTGGGCCATTGCTGGCATTGTGATTGGA AsnAlaLeuProGlnGluAsnGlyLeuSerProGlyAlaIleAlaGlyIleValIleGly	

35 1390 1410 1430 . . .
 GTAGTGGCCCTGGTTGCTCTGATAGCAGTAGCCCTGGCATGTTTCTGCATTCGGGAAG
 ValValAlaLeuValAlaLeuIleAlaValAlaLeuAlaCysPheLeuHisPheGlyLys

40 1450 1470 1490
 ACCGGCAGGGCAAGCGACCAGCGTGTCTCACAGAGCACAAACCTCAGTCTCCAACCA
 ThrGlyArgAlaSerAspGlnArgAspLeuThrGluHisLysProSerValSerAsnHis

45 1510 1530 1550
 A C T C A G G A C C A C T C C A A T G A C C C A C C T A A C A A G A T G A A G T T A C T T A T T C T A C C C T G
 Thr Glu Asp His Ser Asn Asp Pro Pro Asn Lys Met Asn Glu Val Thr Tyr Ser Thr Leu

50 1570 1590 1610

AACTTGAAGCCCAGCAACCCACACAACCAACTTCAGCCTCCCCATCCCTAACAGCCACA
AsnPheGluAlaGlnGlnProThrGlnProThrSerAlaSerProSerLeuThrAlaThr

5

1630 1650 1670

GAAATAATTATTCAAGTAAAAAGCAGTAATGAAACCTGTCCTGCTCACTGCAGTGC
GluIleIleTyrSerGluValLysGln

10

1690 1710 1730

TGATGTATTCAAGTCTCTCACCCCTCATCACTAGGAGATTCCCTTCCCTGTAGGGTAGA

15

1750 1770 1790

GGGGTGGGGACAGAAACAACCTTCTCCTACTCTTCCTCCTAATAGGCATCTCCAGGCTG

20

1810 1830 1850

CCTGGTCACTGCCCTCTCAGTGTCAATAGATGAAAGTACATTGGAGTCTGTAGGAA

25

1870 1890 1910

ACCCAACCTTCTTGTCAATTGAAATTGGCAAAGCTGACTTGGAAAGAGGGACCAGAAC

30

1930 1950 1970

TTCCCCCTCCCTTCCCCCTTCCAAACCTGGACTTGTAAACTTGCCCTGTTCAGAGCAC

35

1990 2010 2030

TCATTCTTCCCACCCCCAGTCCTGCTATCACTCTAATTGGATTGCCATAGCTTG

40

2050 2070 2090

AGGTTATGTCCTTTCCATTAAAGTACATGTGCCAGGAAACAGCGAGAGAGAGAAAGTAAA

45

2110 2130 2150

CGGCAGTAATGCTTCTCCTATTCTCCAAAGCCTTGTGTGAACTAGCAAAGAGAAGAAAA

55

2170 2190 2210

TCAAATATATAACCAATAGTGAAATGCCACAGGTTGTCCACTGTCAGGGTTGTCTACCT

2230 2250 2270
 GTAGGATCAGGGTCTAACGCACCTTGGTGCTTAGCTAGAATAACCACCTAACCTCTGGCA
 5
 2290 2310 2330
 AGCCTGTCTTCAGAGAACCCACTAGAAGCAACTAGGAAAAATCACTTGCCAAAATCCAAG
 10
 2350 2370 2390
 GCAATTCTGATGGAAAATGCAAAAGCACATATATGTTTAATATCTTATGGGCTCTGT
 15
 2410 2430 2450
 TCAAGGCAGTGCTGAGAGGGAGGGGTTATAGCTTCAGGAGGGAACCAGCTCTGATAAAC
 20
 2470 2490 2510
 ACAATCTGCTAGGAACCTGGAAAGGAATCAGAGAGCTGCCTTCAGCGATTATTTAAAT
 25
 2530 2550 2570
 TGTTAAAGAATACACAATTTGGGTATTGGGATTTCTCCTTCTCTGAGACATTCCA
 30
 2590 2610 2630
 CCATTTAATTTGTAACTGCTATTTATGTGAAAAGGGTTATTTACTTAGCTTAGC
 35
 2650 2670 2690
 TATGTCAGCCAATCCGATTGCCTTAGGTGAAAGAAACCACCGAAATCCCTCAGGTCCCTT
 40
 2710 2730 2750
 GGTCAGGAGCCTCTCAAGATTTTGTAGAGCTCCAAATAGAAAATAAGAAAAGGT
 45
 2770 2790 2810
 TTTCTTCATTCATGGCTAGAGCTAGATTTACTCAGTTCTAGGCACCTCAGACCAATCA
 50
 2830 2850 2870
 TCAACTACCATTCTATTCCATGTTGCACCTGTGCATTTCTGTTGCCCCATTCACTT

2890	2910	2930
TGTCAAGGAAACCTGGCCTTGCTAAGGTGTATTTGGTCCITGAGAAGTGGGAGCACCC		
5		
2950	2970	2990
ACACGGGACACTATCACTCATGCTGGTGGCATTGTTACAGCTAGAAAGCTGCACGGTGC		
10		
3010	3030	3050
TAATGCCCTGGAAATGGGGCTGTGAGGAGGAGGATTATAACTTAGGCCTAGCCTCTT		
15		
3070	3090	3110
TTAACAGCCCTCTGAAATTTATCTTTCTTCTATGGGTCTATAAAATGTATCTTATAATAA		
20		
3130	3150	3170
AAAGGAAGGACAGGAGGAAGACAGGCAAATGTACTTCTCACCCAGTCTTCTACACAGATG		
25		
3190	3210	3230
GAATCTCTTGGGCTAAGAGAAAGGTTTATTCTATATTGCTTACCTGATCTCATGTTA		
30		
3250	3270	3290
GGCCTAAGAGGCTTCCTCCAGGAGGATTAGCTTGGAGTTCTCTATAACTCAGGTACCTCTT		
35		
3310	3330	3350
TCAGGGTTTCTAACCTGACACGGACTGTGCATACTTCCCTCATCCATGCTGTGCTGT		
40		
3370	3390	3410
CTTATTTAATTTTCTGGCTAAGATCATGTCTGAATTATGTATGAAAATTATCTATGT		
45		
3430	3450	
TTTTATAATAAAATAATATATCAGACATCGAAAAAA		

(d)

10 20 30 40 50

5 CC 666 66A CAC 6CA 666 CCA ACA 61C ACA 6CA 6CC 616 ACC 6CA 6CA 61C 616 66G 61C

60 70 80 90 100 110

10 AAG 61C TCT ACA AAG 666 ACA 6A6 AAG ACA 6CA 6A6 ACC 61G 66A CCC 6CC 6CA

15 10 Met Gly Pro Pro Ser

120 130 140 150 160 170

15 GCC CCT CCC T6C AGA 716 CAT 6TC CCC 766 AAG 6A6 6TC 716 61C ACA 6CC 6CA 677

20 15 Ala Pro Pro Cys Arg Leu His Val Pro Trp Lys Glu Val Leu Leu Thr Ala Ser Leu

180 190 200 210 220 230

20 CTA ACC 61C 766 AAG CCA 6CC ACC ACT 6CC 6A6 6TC ACT ATT 6AA 61C 6CA 61C

25 20 Leu Thr Phe Trp Asn Pro Pro Thr Thr Ala Lys Leu Thr Ile Glu Ser Thr Pro Phe

240 250 260 270 280

25 240 250 260 270 280

25 AAT 6TC 6CA 6A6 666 6A6 6A6 677 C17 CTA 61C 6CC CAC 6AC 616 CCC 6A6 AAT 661

30 25 Asn Val Ala Glu Glu Lys Glu Val Leu Leu Leu Ala His Asn Leu Pro Glu Asn Arg

30 25 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28

290 300 310 320 330 340

30 ATT 661 TAC AGC 666 TAC AAA 666 CAA 666 GAA 666 G16 6A6 666 GGC 666 GAA 666

35 30 Ile Glu Tyr Ser Trp Tyr Lys Glu Glu Arg Val Asp Glu Asn Ser Leu Ile Val Glu

35 30 27 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47

35 350 360 370 380 390 400

35 TAT 67A ATA 66A ACT CAA CAA 667 ACC CAA 666 CCC 6CA TAC AGT 667 C6A 6A6 ACA

40 35 Tyr Val Ile Glu Ile Glu Glu Ala Thr Pro Glu Pro Ala Tyr Ser Glu Arg Glu Thr

40 35 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66

410 420 430 440 450

40 ATA TAC CCC 6A6 6CA 6CC 616 616 6TC 6TC 6TC 6AC 61C ACC 6A6 AAT 6AC 6CA 66A 61C

45 40 Ile Tyr Pro Asn Ala Ser Leu Leu Ile Glu Asn Val Thr Glu Asn Asp Thr Glu Phe

45 40 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85

460 470 480 490 500 510

45 TAC ACC CTA CAA 61C ATA AAG TCA 6A6 CTT 676 AAT 6A6 6A6 6CA ACC 66A 6A6 61C

50 45 Tyr Ile Leu Glu Val Ile Lys Ser Asp Leu Val Asn Glu Glu Ala Thr Glu Glu Phe

50 45 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87

520 530 540 550 560 570

50 CAT 67A TAC CCC 6A6 616 CCC 6A6 CCC 6CC TCC 6TC 6TC AGC 6AC 6AC 6CC 616

50 His Val Tyr Pro Glu Leu Pro Ser Ile Ser Ser Asn Asn Ser Asn Pro Val

50 50 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123

EP 0 346 710 B1

EP 0 346 710 B1

1150 1160 1170 1180 1190
 T66 TGT ATT TTC GAT ATT TCA GGA AGA CTC GCA GAT T66 ACC AGA CCC TGA ATT CTT

5 1200 1210 1220 1230 1240 1250
 CTA GCT CCT CCA ATC CCA TTT TAT CCC ATG GAA CCA CTA AAA ACA AGG TCT GCT CTC

10 1260 1270 1280 1290 1300 1310
 CTC CTC AAG CCC TAT ATG CTC GAG ATG GAC AAC TCA ATG AAA ATT TAA AGG AAA AAC

15 1320 1330 1340 1350 1360 1370
 CCT CAG GCG TGA GGT GTC CAC TCA GAG ACT TCA CCT AAC TAA AGA CAG GCA AAC

20 1380 1390 1400 1410 1420
 TGC AAA CCA AAC CTC TTT CGC TTC GCA GGA TGA TGG TGT CAT TAG TAT TTC ACA AGA

25 1430 1440 1450 1460 1470 1480
 GGT AGC TTC AGA GGG TAA CTT AAC AGA GTC TCA GAT CTA CCT TGT CAA TCC CAA GGT

30 1490 1500 1510 1520 1530 1540
 TTT ACA TAA ATT AAG CGA TCC TTT AGT GCA CCC AGT GAC TGA CAT TAG CAG CAT CTT

35 1550 1560 1570 1580 1590
 TAA CAC AGC CTC GTC TTC AAG TGT ACA GTC CTT TTC AGA GTT GGA AAC ACT CCA

40 1600 1610 1620 1630 1640 1650
 ACT GAA ATG TTA AGG AAG AAC ATA GAT CCA ATT AAA AAA ATT TAA AAC CAA TTT AAA

45 1660 1670 1680 1690 1700 1710
 AGA AAA AAA GAA AAC AGG AGA TTC CAG TCT ACT TGA GTT ABC ATA ATA CAG AAG TCC

50 1720 1730 1740 1750 1760
 CCT CTA CTT TAA CTT TTA CAA AAA AGT AAC CTC AAC TAA TCT GAT GTT AAC CAA TGT

1770 1780 1790 1800 1810 1820
 ATT TAT T16 TCT GGT TCT 611 TCC T16 T1C CAA T1T GAC AAA ACC CAC T1T TCT TCT

5

1830 1840 1850 1860 1870 1880
 ATT G1A T1G CCC AGG GGG AGC TAT CAC TGT ACT T1T AGA G1G G1G C1G T1A G1T

10

1890 1900 1910 1920 1930 1940
 CAT AAA TCA CAA ATA AAA GGC AAT TAT C1C TAT AAC TAA AAA AAA AAA AAA AAA AAA

15

1950 1960
 AAA AAA AAA AAA AAA AAA AAA AAA

20

A schematic relationship of the transmembrane CEA's, namely TM-1 (CEA-(c)), TM-2 (CEA-(e)), TM-3 (CEA-(f)) and TM-4 (CEA-(g)) is depicted in Fig. 1:

Assuming TM-1 is composed of five sections as depicted in Fig. 1, namely 10, 12, 14, 16 and 18, TM-2 differs from TM-1 in that the 100 amino acid (100 AA) section 14 is deleted and at splice point 20 between 25 sections 12 and 16, surprisingly an extra amino acid, namely Asp occurs.

TM-3 is the same as TM-1 except that section 18 is truncated at splice point 22, i.e., a section of 70 amino acids is deleted and results in a new section made up of subsections 24 + 26. Surprisingly, however, six new amino acids (section 26) occur. Another example of formation of a novel amino acid sequence resulting from a deletion of nucleic acid sequence is for platelet derived growth factor-A.

30 TM-4 is the same as TM-2 up until the end of subsection 24.

Subsection 24 is contained in section 18 of TM-1 and TM-2, but is not depicted in Fig. 1 for TM-1 and TM-2.

Some CEA epitopes are unique. These are the epitopes which have been useful for distinguishing the various CEA-like antigens immunologically. Peptide epitopes are defined by the linear amino acid sequence 35 of the antigen and/or features resulting from protein folding. The information required for protein folding is encoded in the primary amino acid sequence. Therefore, antigenic differences ultimately result from differences in the primary structure of the different CEA molecules. The differences residing in the CEA protein in the CEA species can thus be determined by determining the primary amino acid sequences. This can be most readily accomplished by cloning and sequencing each of the genes for CEA. To determine 40 which gene products will be most useful for cancer diagnosis, unique probes can be selected for each gene and expression of each gene can be determined in different tumor types by nucleic acid hybridization techniques. The present invention provides a tool with which to identify potential genes coding for different members of the CEA family and to determine the theoretical primary amino acid sequences for them. Using the method of automated peptide synthesis, peptides can then be synthesized corresponding to unique 45 sequences in these antigens. With these peptides, antibodies to these sequences can be produced which, in the intact CEA molecule, might not be recognized by the animal being immunized. Having accomplished this, advantage can then be taken of the differences in these antigens to generate specific immunoassays for the measurement of each antigen.

A wide variety of host/cloning vehicle combinations may be employed in cloning the double-stranded 50 nucleic acid prepared in accordance with this invention. For example, useful cloning vehicles may consist of segments of chromosomal, non-chromosomal and synthetic DNA sequences, such as various known derivatives of SV40 and known bacterial plasmids, e.g., plasmids from E. coli including col E1, pCR1, pBR322, pMB89 and their derivatives, wider host range plasmids, e.g., RP4, and phage DNAs, e.g., the numerous derivatives of phage, e.g., NM989, and other DNA phages, e.g., M13 and Filamentous single- 55 stranded DNA phages and vectors derived from combinations of plasmids and phage DNAs such as plasmids which have been modified to employ phage DNA or other expression control sequences or yeast plasmids such as the 2 μ plasmid or derivatives thereof. Useful hosts may include bacterial hosts such as strains of E. coli, such as E. coli HB 101, E. coli X1776, E. coli X2282, E. coli MRCI and strains of

Pseudomonas, Bacillus subtilis, Bacillus stearothermophilus and other E. coli, bacilli, yeasts and other fungi, animal or plant hosts such as animal (including human) or plant cells in culture or other hosts. Of course, not all host/vector combinations may be equally efficient. The particular selection of host/cloning vehicle combination may be made by those of skill in the art after due consideration of the principles set forth

5 without departing from the scope of this invention.

Furthermore, within each specific cloning vehicle, various sites may be selected for insertion of the nucleic acid according to the present invention. These sites are usually designated by the restriction endonuclease which cuts them. For example, in pBR322 the PstI site is located in the gene for beta-lactamase, between the nucleotide triplets that code for amino acids 181 and 182 of that protein. One of the 10 two HindII endonuclease recognition sites is between the triplets coding for amino acids 101 and 102 and one of the several Taq sites at the triplet coding for amino acid 45 of beta-lactamase in pBR322. In similar fashion, the EcoRI site and the PVUII site in this plasmid lie outside of any coding region, the EcoRI site being located between the genes coding for resistance to tetracycline and ampicillin, respectively. These 15 sites are well recognized by those of skill in the art. It is, of course, to be understood that a cloning vehicle useful in this invention need not have a restriction endonuclease site for insertion of the chosen DNA fragment. Instead, the vehicle could be cut and joined to the fragment by alternative means.

The vector or cloning vehicle and in particular the site chosen therein for attachment of a selected nucleic acid fragment to form a recombinant nucleic acid molecule is determined by a variety of factors, e.g., the number of sites susceptible to a particular restriction enzyme, the size of the protein to be 20 expressed, the susceptibility of the desired protein to proteolytic degradation by host cell enzymes, the contamination of the protein to be expressed by host cell proteins difficult to remove during purification, the expression characteristics, such as the location of start and stop codons relative to the vector sequences, and other factors recognized by those of skill in the art. The choice of a vector and an insertion site for a 25 particular gene is determined by a balance of these factors, not all sections being equally effective for a given case.

Methods of inserting nucleic acid sequences into cloning vehicles to form recombinant nucleic acid molecules include, for example, dA-dT tailing, direct ligation, synthetic linkers, exonuclease and polymerase-linked repair reactions followed by ligation, or extension of the nucleic acid strand with an appropriate polymerase and an appropriate single-stranded template followed by ligation.

30 It should also be understood that the nucleotide sequences or nucleic acid fragments inserted at the selected site of the cloning vehicle may include nucleotides which are not part of the actual structural gene for the desired polypeptide or mature protein or may include only a fragment of the complete structural gene for the desired protein or mature protein.

The cloning vehicle or vector containing the foreign gene is employed to transform an appropriate host 35 so as to permit that host to replicate the foreign gene and to express the protein coded by the foreign gene or portion thereof. The selection of an appropriate host is also controlled by a number of factors recognized by the art. These include, for example, the compatibility with the chosen vector, the toxicity of proteins encoded by the hybrid plasmid, the ease of recovery of the desired protein, the expression characteristics, biosafety and costs. A balance of these factors must be struck with the understanding that not all hosts may 40 be equally effective for expression of a particular recombinant DNA molecule.

The level of production of a protein is governed by two major factors: the number of copies of its gene within the cell and the efficiency with which those gene copies are transcribed and translated. Efficiency of transcription and translation (which together comprise expression) is in turn dependent upon nucleotide sequences, normally situated ahead of the desired coding sequence. These nucleotide sequences or 45 expression control sequences define *inter alia*, the location at which RNA polymerase interacts to initiate transcription (the promoter sequence) and at which ribosomes bind and interact with the mRNA (the product of transcription) to initiate translation. Not all such expression control sequences function with equal efficiency. It is thus of advantage to separate the specific coding sequences for the desired protein from their adjacent nucleotide sequences and fuse them instead to other known expression control sequences so 50 as to favor higher levels of expression. This having been achieved, the newly engineered nucleic acid, e.g., DNA, fragment may be inserted into a multicopy plasmid or a bacteriophage derivative in order to increase the number of gene copies within the cell and thereby further improve the yield of expressed protein.

Several expression control sequences may be employed as described above. These include the operator, promoter and ribosome binding and interaction sequences (including sequences such as the 55 Shine-Dalgarno sequences) of the lactose operon of E. coli ("the lac system"), the corresponding sequences of the tryptophan synthetase system of E. coli ("the trp system"), the major operator and promoter regions of phage λ ($O_L P_L$ and $O_R P_R'$), the control region of Filamentous single-stranded DNA phages, or other sequences which control the expression of genes of prokaryotic or eukaryotic cells and

their viruses. Therefore, to improve the production of a particular polypeptide in an appropriate host, the gene coding for that polypeptide may be selected and removed from a recombinant nucleic acid molecule containing it and reinserted into a recombinant nucleic acid molecule closer or in a more appropriate relationship to its former expression control sequence or under the control of one of the above described expression control sequences. Such methods are known in the art.

As used herein "relationship" may encompass many factors, e.g., the distance separating the expression enhancing and promoting regions of the recombinant nucleic acid molecule and the inserted nucleic acid sequence, the transcription and translation characteristics of the inserted nucleic acid sequence or other sequences in the vector itself, the particular nucleotide sequence of the inserted nucleic acid sequence and other sequences of the vector and the particular characteristics of the expression enhancing and promoting regions of the vector.

Further increases in the cellular yield of the desired products depend upon an increase in the number of genes that can be utilized in the cell. This is achieved, for illustration purposes, by insertion of recombinant nucleic acid molecules engineered into the temperate bacteriophage λ (NM989), most simply by digestion of the plasmid with a restriction enzyme, to give a linear molecule which is then mixed with a restricted phage λ cloning vehicle (e.g., of the type described by N. E. Murray et al, "Lambda Phages That Simplify the Recovery of In Vitro Recombinants", *Molec. Gen. Genet.*, 150, pp. 53-61 (1977) and N. E. Murray et al, "Molecular Cloning of the DNA Ligase Gene From Bacteriophage T4", *J. Mol. Biol.*, 132, pp. 493-505 (1979)) and the recombinant DNA molecule recircularized by incubation with DNA ligase. The desired recombinant phage is then selected as before and used to lysogenize a host strain of *E. coli*.

Particularly useful λ cloning vehicles contain a temperature-sensitive mutation in the repression gene *cl* and suppressible mutations in gene *S*, the product of which is necessary for lysis of the host cell, and gene *E*, the product of which is major capsid protein of the virus. With this system, the lysogenic cells are grown at 32°C and then heated to 45°C to induce excision of the prophage. Prolonged growth at 37°C leads to high levels of production of the protein, which is retained within the cells, since these are not lysed by phage gene products in the normal way, and since the phage gene insert is not encapsulated it remains available for further transcription. Artificial lysis of the cells then releases the desired product in high yield.

In addition, it should be understood that the yield of polypeptides prepared in accordance with this invention may also be improved by substituting different codons for some or all of the codons of the present DNA sequences, these substituted codons coding for amino acids identical to those coded for by the codons replaced.

Finally, the activity of the polypeptides produced by the recombinant nucleic acid molecules of this invention may be improved by fragmenting, modifying or derivatizing the nucleic acid sequences or polypeptides of this invention by well-known means, without departing from the scope of this invention.

The polypeptides of the present invention include the following:

- (1) the polypeptides expressed by the above described cells,
- (2) polypeptides prepared by synthetic means,
- (3) fragments of polypeptides (1) or (2) above, such fragments produced by synthesis of amino acids or by digestion or cleavage.

Regarding the synthetic peptides according to the invention, chemical synthesis of peptides is described in the following publications: S.B.H. Kent, *Biomedical Polymers*, eds. Goldberg, E.P. and Nakajima, A. (Academic Press, New York), 213-242, (1980); A.R. Mitchell, S.B.H. Kent, M. Engelhard and R.B. Merrifield, *J. Org. Chem.*, 43, 2845-2852, (1978); J.P. Tam, T.-W. Wong, M. Riemen, F.-S. Tjoeng and R.B. Merrifield, *Tet. Letters*, 4033-4036, (1979); S. Mojsov, A.R. Mitchell and R.B. Merrifield, *J. Org. Chem.*, 45, 555-560, (1980); J.P. Tam, R.D. DiMarchi and R.B. Merrifield, *Tet. Letters*, 2851-2854, (1981); and S.B.H. Kent, M. Riemen, M. Le Doux and R.B. Merrifield, *Proceedings of the IV International Symposium on Methods of Protein Sequence Analysis*, (Brookhaven Press, Brookhaven, NY), in press, 1981.

In the Merrifield solid phase procedure, the appropriate sequence of L-amino acids is built up from the carboxyl terminal amino acid to the amino terminal amino acid. Starting with the appropriate carboxyl terminal amino acid attached to a polystyrene (or other appropriate) resin via chemical linkage to a chloromethyl group, benzhydrylamine group, or other reactive group of the resin, amino acids are added one by one using the following procedure. The peptide-resin is:

- (a) washed with methylene chloride;
- (b) neutralized by making for 10 minutes at room temperature with 5% (v/v) diisopropylethylamine (or other hindered base) in methylene chloride;
- (c) washed with methylene chloride;
- (d) an amount of amino acid equal to six times the molar amount of the growing peptide chain is activated by combining it with one-half as many moles of a carbodiimide (e.g., dicyclohexylcarbodiimide,

or diisopropylcarbodiimide) for ten minutes at 0 °C, to form the symmetric anhydride of the amino acid. The amino acid used should be provided originally as the N-alpha-tert.-butyloxycarbonyl derivative, with side chains protected with benzyl esters (e.g., aspartic or glutamic acids), benzyl ethers (e.g., serine, threonine, cysteine or tyrosine), benzyloxycarbonyl groups (e.g., lysine) or other protecting groups commonly used in peptide synthesis;

- 5 (e) the activated amino acid is reacted with the peptide-resin for two hours at room temperature, resulting in addition of the new amino acid to the end of the growing peptide chain;
- (f) the peptide-resin is washed with methylene chloride;
- (g) the N-alpha-(tert.-butyloxycarbonyl) group is removed from the most recently added amino acid by reacting with 30 to 65%, preferably 50% (v/v) trifluoroacetic acid in methylene chloride for 10 to 30 minutes at room temperature;
- (h) the peptide-resin is washed with methylene chloride;
- (i) steps (a) through (h) are repeated until the required peptide sequence has been constructed.

The peptide is then removed from the resin and simultaneously the side-chain protecting groups are removed, by reaction with anhydrous hydrofluoric acid containing 10% v/v of anisole or other suitable (aromatic) scavenger. Subsequently, the peptide can be purified by gel filtration, ion exchange, high pressure liquid chromatography, or other suitable means.

In some cases, chemical synthesis can be carried out without the solid phase resin, in which case the synthetic reactions are performed entirely in solution. The reactions are similar and well known in the art, and the final product is essentially identical.

Digestion of the polypeptide can be accomplished by using proteolytic enzymes, especially those enzymes whose substrate specificity results in cleavage of the polypeptide at sites immediately adjacent to the desired sequence of amino acids.

Cleavage of the polypeptide can be accomplished by chemical means. Particular bonds between amino acids can be cleaved by reaction with specific reagents. Examples include the following: bonds involving methionine are cleaved by cyanogen bromide; asparaginyl-glycine bonds are cleaved by hydroxylamine.

The present invention has the following advantages:

- (1) The nucleic acids coding for TM-1, TM-2 and TM-3 can be used as probes to isolate other members of the CEA gene family.
- 30 (2) The nucleic acids coding for TM-1, TM-2 and TM-3 can be used to derive oligonucleotide probes to determine the expression of TM-1, TM-2, TM-3 and other CEA genes in various tumor types.
- (3) TM-1, TM-2, TM-3 and TM-4 nucleotide sequences can be used to predict the primary amino acid sequence of the protein for production of synthetic peptides.
- (4) Synthetic peptides derived from the above sequences can be used to produce sequence-specific antibodies.
- 35 (5) Immunoassays for each member of the CEA antigen family can be produced with these sequence-specific antibodies and synthetic peptides.
- (6) The aforementioned immunoassays can be used as diagnostics for different types of cancer if it is determined that different members of the CEA family are clinically useful markers for different types of cancer.

Peptides according to the present invention can be labelled by conventional means using radioactive moieties (e.g., ¹²⁵I), enzymes, dyed and fluorescent compounds, just to name a few.

Several possible configurations for immunoassays according to the present invention can be used. The readout systems capable of being employed in these assays are numerous and non-limiting examples of such systems include fluorescent and colorimetric enzyme systems, radioisotopic labelling and detection and chemiluminescent systems. Two examples of immunoassay methods are as follows:

- (1) An enzyme linked immunoassay (ELISA) using an antibody preparation according to the present invention (including Fab or F(ab)' fragments derived therefrom) to a solid phase (such as a microtiter plate or latex beads) is attached a purified antibody of a specificity other than that which is conjugated to the enzyme. This solid phase antibody is contacted with the sample containing CEA antigen family members. After washing, the solid phase antibody-antigen complex is contacted with the conjugated anti-peptide antibody (or conjugated fragment). After washing away unbound conjugate, color or fluorescence is developed by adding a chromogenic or fluorogenic substrate for the enzyme. The amount of color or fluorescence developed is proportional to the amount of antigen in the sample.
- 50 (2) A competitive fluorometric immunoassay using fluorescently labelled peptide or synthetic peptides of the sequences for TM-2, TM-2, TM-3 and TM-4. In this example, the purified peptide expressed by cells or synthetic peptides thereof are fluorescently labelled. To a solid phase is attached a purified antibody. This solid phase is then contacted with sample containing CEA antigen family members to which has

been added fluorescent peptide probe. After binding, excess probe is washed away the amount of bound probe is quantitated. The amount of bound fluorescent probe will be inversely proportional to the amount of antigen in the sample.

In the nucleic acid hybridization method according to the present invention, the nucleic acid probe is conjugated with a label, for example, an enzyme, a fluorophore, a radioisotope, a chemiluminescent compound, etc. In the most general case, the probe would be contacted with the sample and the presence of any hybridizable nucleic acid sequence would be detected by developing in the presence of a chromogenic enzyme substrate, detection of the fluorophore by epifluorescence, by autoradiography of the radioisotopically labelled probe or by chemiluminescence. The detection of hybridizable RNA sequences can be accomplished by (1) a dot blot methodology or (2) an *in situ* hybridization methodology. Methods for these last two techniques are described by D. Gillespie and J. Bresser, "mRNA Immobilization in *Nal*: Quick Blots", *Biotechniques*, 184-192, November/December 1983 and J. Lawrence and R. Singer, "Intracellular Localization of Messenger RNAs for Cytoskeletal Proteins", *Cell*, 45, 407-415, May 9, 1986, respectively. The readout systems can be the same as described above, e.g., enzyme labelling, radiolabelling, etc.

As stated above, the invention also relates to the use in medicine of the aforementioned complex of the invention.

The invention further provides a pharmaceutical composition containing as an active ingredient a complex of the invention in the form of a sterile and/or physiologically isotonic aqueous solution.

For parenteral administration, solutions and emulsions containing as an active ingredient the complex of the invention should be sterile and, if appropriate, blood-isotonic.

It is envisaged that the active complex will be administered perorally, parenterally (for example, intramuscularly, intraperitoneally, or intravenously), rectally or locally.

Example 1: Preparation of cDNA in pcE22 which codes for TM2-CEA [CEA-(e)]

25 Example 1a: RNA Preparation

Messenger RNA was prepared by the proteinase K extraction method of J. Favolaro, R. Treisman and R. Kamen, *Methods in Enzymology*, 65, 718, Academic Press, Inc. (1980), followed by oligo dT cellulose chromatography to yield poly A+ RNA (3'-polyadenylated eukaryotic RNA containing most mRNA sequences that can be translated into polypeptides). To obtain approximately 100 µg of poly A+ RNA, approximately 3×10^8 cells of transfected 23.411 (ATCC No. CRL 9731, deposited with the ATCC on June 1, 1988), that expresses TM-1, TM-2, TM-3 and TM-4, Kamarck et al, *Proc. Natl. Acad. Sci., USA*, 84, 5350-5354, August 1987, were harvested from roller bottles after late logarithmic growth. Cells were lysed by homogenization in an ice-cold solution of 140 mM NaCl, 1.5 mM MgCl₂, 10 mM Tris-HCl, pH 8.0, 0.5% NP40®, 4 mM dithiothreitol and 20 units of placental ribonuclease inhibitor/ml. Sodium deoxycholate was then added to 0.2%. Cytoplasm and nuclei were separated by centrifugation of the homogenate at 12,000xg for 20 minutes. The cytoplasmic fraction was mixed with an equal volume of 0.2 M Tris-HCl, pH 7.8, 25 mM EDTA, 0.3 M NaCl, 2% sodium dodecyl sulfate and 400 µg/ml of proteinase K, incubated for 1 hour at 37°C, then extracted once with an equal volume of phenol/chloroform (1:1/v:v) solution. Nucleic acids were obtained by ethanol precipitation of the separated aqueous phase. Total RNA was enriched by passage in 0.5 M NaCl, 10 mM Tris-HCl, pH 7.8, 0.1% sarcosyl® through an oligo dT(12-18) cellulose column. After washing, bound RNA was eluted in the same solution without sodium chloride.

45 Example 1b: Reverse Transcription of mRNA

Ten micrograms of poly A+ RNA were primed for reverse transcription with oligo dT(12-18) and pdN₆ primers. One hundred microliter reaction was performed for 4 hours at 42°C with 200 units AMV reverse transcriptase (Life Science, Inc. St. Petersburg, Florida, U.S.A.). The RNA component of the cDNA/mRNA hybrids was replaced with the second complementary strand by treatment with RNase H, *E. coli* DNA polymerase I and *E. coli* DNA ligase at 12°C and 22°C for 1.5 hours each. Molecular ends were polished by treatment with T4 DNA polymerase. cDNA was phenol/chloroform extracted and purified over a "SEPHADEX® G-50" spun column prepared in 10 mM Tris-HCl, pH 7.8, 1 mM EDTA (TE).

55 Example 1c: Cloning of pcE22 (plasmid cDNA E22)

Synthetic DNA linkers 5' pCCCGGG 3'
 3' GGGCCCTTAA 5'

were attached to the ends of cDNA by blunt end ligation with excess T4 DNA ligase. Excess linkers were removed by chromatography through "SEPHADEX® G-50" (medium) in TE, and by fractionation on 0.8% low melting agarose gel. Based on Northern blot analysis of poly A+ RNA of the 23.411 cell line, the size of the CEA-related mRNA was estimated at 3.6 kb. Therefore, cDNA fragments between 2 and 4 kb were recovered from gel slices and fragments were ethanol precipitated. After resuspension of cDNA in TE, EcoRI-cleaved lambda gt10 arms were added to cDNA at an estimated molar ratio of 1:1. Ligation proceeded at 7 °C for 2 days in the presence of T4 DNA ligase. Aliquots of the ligation reaction were added to commercially-obtained packaging mix (Stratagene, San Diego, California, U.S.A.). Five million phage particles were obtained after in vitro packaging and infection of E. coli host NM514.

10

Example 1d: Screening of Recombinant Library

Five hundred thousand packaged lambda particles were plated on lawns of E. coli NM514 and replicate patterns were lifted onto nitrocellulose sheets as described by W.D. Benton and R.W. Davis, Science 196, 180-182, (1977). Positive phage were selected by hybridization with ³²P-labeled LV7 cDNA insert probe that contained a domain repeated among various CEA family members. By multiple rounds of screening. Phage from individual plaques were amplified and titered, and these were used to prepare small quantities of recombinant phage DNA.

20

Example 1e: DNA Manipulation

Phage DNA was prepared according to T. Maniatis, E. Fritsch and J. Sambrook, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor, (1982). DNA segments were isolated from low melting agarose gels and inserted for subcloning into Bluescript plasmid vectors (Stratagene, San Diego, California, U.S.A.). DNA sequencing was performed by the dideoxy termination method of F. Sanger, S. Nicklen and A. Coulson, Proc. Natl. Acad. Sci., U.S.A., 74, 5463-5467, (1977). The nucleic acid and translated sequence for cDNA in pcE22 is given hereinabove (TM-2 (CEA-(e))).

30

Example 2: Preparation of cDNA in pcHT-6 which Partically Codes for TM3-CEA [CEA-(f)]

Example 2a: RNA Preparation

Messenger RNA was prepared by the proteinase K extraction method of J. Favolaro, R. Treisman and R. Kamen, Methods in Enzymology, 65 718, Academic Press, Inc. (1980), followed by oligo dT cellulose chromatography to yield poly A+ RNA (3'-polyadenylated eukaryotic RNA containing most mRNA sequences that can be translated into polypeptides). To obtain approximately 100 ug of poly A+ RNA, approximately 3×10^8 cells of HT-29 tumor cells (ATCC HTB38) were harvested from roller bottles after late logarithmic growth. Cells were lysed by homogenization in an ice-cold solution of 140 mM NaCl, 1.5 mM MgCl₂, 10 mM Tris-HCl, pH 8.0, 0.5% NP40®, 4 mM dithiothreitol and 20 units of placental ribonuclease inhibitor/ml. Sodium deoxycholate was then added to 0.2%. Cytoplasm and nuclei were separated by centrifugation of the homogenate at 12,000xg for 20 minutes. The cytoplasmic fraction was mixed with an equal volume of 0.2 M Tris-HCl, pH 7.8, 25 mM EDTA, 0.3 M NaCl, 2% sodium dodecyl sulfate and 400 µg/ml of proteinase K, incubated for 1 hour at 37 °C, then extracted once with an equal volume of phenol/chloroform (1:1/v:v) solution. Nucleic acids were obtained by ethanol precipitation of the separated aqueous phase. Total RNA was enriched by passage in 0.5 M NaCl, 10 mM Tris-HCl, pH 7.8, 0.1% sarcosyl® through an oligo dT(12-18) cellulose column. After washing, bound RNA was eluted in the same solution without sodium chloride.

50

Example 2b: Reverse Transcription of mRNA

Ten micrograms of HT-29 poly A+ RNA were primed for reverse transcription with oligo dT(12-18) and pdN₆ primers. One hundred microliter reaction was performed for 4 hours at 42 °C with 200 units AMV reverse transcriptase (Life Science, Inc. St. Petersburg, Florida, U.S.A.). The RNA component of the cDNA/mRNA hybrids was replaced with the second complementary strand by treatment with RNase H, E. coli DNA polymerase I and E. coli DNA ligase at 12 °C and 22 °C for 1.5 hours each. Molecular ends were polished by treatment with T4 DNA polymerase. cDNA was phenol/chloroform extracted and purified over a "SEPHADEX® G-50" spun column prepared in 10 mM Tris-HCl, pH 7.8, 1 mM EDTA (TE).

Example 2c: Cloning of pcHT-6 (plasmid cDNA HT-6)

Synthetic DNA linkers 5' pCCCGGG 3'
3' GGGCCCTTAA 5'

5 were attached to the ends of cDNA by blunt end ligation with excess T4 DNA ligase. Excess linkers were removed by chromatography through "SEPHADEX® G-50" (medium) in TE, and by fractionation on 0.8% low melting agarose gel. Based on Northern blot analysis of poly A+ RNA of the HT-29 cell line, the size of the CEA-related mRNA was estimated at 2.2 kb. Therefore, cDNA fragments between 2 and 3 kb were recovered from gel slices and fragments were ethanol precipitated. After resuspension of cDNA in TE,
 10 EcoRI-cleaved lambda gt10 arms were added to cDNA at an estimated molar ratio of 1:1. Ligation proceeded at 7 °C for 2 days in the presence of T4 DNA ligase. Aliquots of the ligation reaction were added to commercially-obtained packaging mix (Stratagene, San Diego, California, U.S.A.). Five million phage particles were obtained after *in vitro* packaging and infection of *E. coli* host NM514.

15 Example 2d: Screening of Recombinant Library

Five hundred thousand packaged lambda particles were plated on lawns of *E. coli* NM514 and replicate patterns were lifted onto nitrocellulose sheets as described by W.D. Benton and R.W. Davis, *Science*, 196, 180-182, (1977). Positive phage were selected by hybridization with ³²P-labeled LV7 cDNA insert probe that contained a domain repeated among various CEA family members. By multiple rounds of screening, Phage from individual plaques were amplified and titered, and these were used to prepare small quantities of recombinant phage DNA.

Example 2e: DNA Manipulation

25 Phage DNA was prepared according to T. Maniatis, E. Fritsch and J. Sambrook, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor, (1982). DNA segments were isolated from low melting agarose gels and inserted for subcloning into Bluescript plasmid vectors (Stratagene, San Diego, California, U.S.A.). DNA sequencing was performed by the dideoxy termination method of F. Sanger, S. Nicklen and A. Coulson, 30 Proc. Natl. Acad. Sci., U.S.A., 74, 5463-5467, (1977). The nucleic acid and translated sequence for cDNA in HT-6 not complete at the 5' end of its coding region, but the nucleotide sequence and restriction map of the HT-6 insert indicates that it is related to nucleic acid sequences of cDNA clones coding for CEA-(c) and CEA-(e). The nucleotide sequence of HT-6 insert differs from these clones at only nucleotide position 1463 to 35 1515 and 1676 to 2429 of the CEA-(c) cDNA. It is inferred from this result that the pcHT-6 insert is a partial coding sequence for CEA-(f) and the presumed nucleic acid and translated sequence of CEA-(f) is given hereinabove (TM-3 (CEA-(f)).

Example 3: Preparation of cDNA which codes for TM4-CEA [CEA-(g)]

40 Example 3a: RNA Preparation

Messenger RNA is prepared by the proteinase K extraction method of J. Favolaro, R. Treisman and R. Kamen, Methods in Enzymology, 65, 718, Academic Press, Inc. (1980), followed by oligo dT cellulose chromatography to yield poly A+ RNA (3'-polyadenylated eukaryotic RNA containing most mRNA sequences that can be translated into polypeptides). To obtain approximately 100 ug of poly A+ RNA, approximately 3×10^8 cells of transfectant 23.411 or tumor cell line HT-29 (ATCC HTB 38) are harvested from roller bottles after late logarithmic growth. Cells are lysed by homogenization in an ice-cold solution of 140 mM NaCl, 1.5 mM MgCl₂, 10 mM Tris-HCl, pH 8.0, 0.5% NP40®, 4 mM dithiothreitol and 20 units of placental ribonuclease inhibitor/ml. Sodium deoxycholate was then added to 0.2%. Cytoplasm and nuclei are separated by centrifugation of the homogenate at 12,000xg for 20 minutes. The cytoplasmic fraction is mixed with an equal volume of 0.2 M Tris-HCl, pH 7.8, 25 mM EDTA, 0.3 M NaCl, 2% sodium dodecyl sulfate and 400 µg/ml of proteinase K, incubated for 1 hour at 37°C, then extracted once with an equal volume of phenol/chloroform (1:1/v:v) solution. Nucleic acids are obtained by ethanol precipitation of the separated aqueous phase. Total RNA is enriched by passage in 0.5 M NaCl, 10 mM Tris-HCl, pH 7.8, 0.1% sarcosyl through an oligo dT(12-18) cellulose column. After washing, bound RNA is eluted in the same solution without sodium chloride.

Example 3b: Reverse Transcription of mRNA

Ten micrograms of 23.411 or HT 29 poly A + RNA are primed for reverse transcription with oligo dT(12-18) and pdN₆ primers. One hundred microliter reaction was performed for 4 hours at 42 °C with 200 units 5 AMV reverse transcriptase (Life Science, Inc. St. Petersburg, Florida, U.S.A.). The RNA component of the cDNA/mRNA hybrids is replaced with the second complementary strand by treatment with RNase H, *E. coli* DNA polymerase I and *E. coli* DNA ligase at 12 °C and 22 °C for 1.5 hours each. Molecular ends are polished by treatment with T4 DNA polymerase. cDNA is phenol/chloroform extracted and purified over a "SEPHADEX® G-50" spun column prepared in 10 mM Tris-HCl, pH 7.8, 1 mM EDTA (TE).

10

Example 3c: Cloning of cDNA for CEA-(g)

Synthetic DNA linkers 5' pCCCGGG 3'
3' GGGCCCTTAA 5'
15 are attached to the ends of cDNA by blunt end ligation with excess T4 DNA ligase. Excess linkers are removed by chromatography through "SEPHADEX® G-50" (medium) in TE, and by fractionation on 0.8% low melting agarose gel. Based on Northern blot analysis of poly A + RNA of the 23.411 and HT-29 cell lines, the size of the CEA-related mRNA is estimated at 1.7 kb. Therefore, cDNA fragments between 1 and 2 kb are recovered from gel slices and fragments are ethanol precipitated. After resuspension of cDNA in 20 TE, EcoRI-cleaved lambda gt10 arms are added to cDNA at an estimated molar ratio of 1:1. Ligation proceeds at 7 °C for 2 days in the presence of T4 DNA ligase. Aliquots of the ligation reaction are added to commercially-obtained packaging mix (Stratagene, San Diego, California, U.S.A.). Phage particles are obtained after *in vitro* packaging and infection of *E. coli* host NM514.

25 Example 3d: Screening of Recombinant Library

Five hundred thousand to one million packaged lambda particles are plated on lawns of *E. coli* NM514 and replicate patterns are lifted onto nitrocellulose sheets as described by W.D. Benton and R.W. Davis, *Science*, 196, 180-182, (1977). Positive phage are selected by hybridization with ³²P-labeled LV7 cDNA 30 insert probe that contained a domain repeated among various CEA family members. By this selection method, positive phage are obtained after multiple rounds of screening. Phage from individual plaques are amplified and titered, and these are used to prepare small quantities of recombinant phage DNA.

Example 3e: DNA Manipulation

35

Phage DNA is prepared according to T. Maniatis, E. Fritsch and J. Sambrook, *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor, (1982). DNA segments are isolated from low melting agarose gels and inserted for subcloning into Bluescript plasmid vectors (Stratagene, San Diego, California, U.S.A.). DNA sequencing is performed by the dideoxy termination method of F. Sanger, S. Nicklen and A. Coulson, *Proc. 40 Natl. Acad. Sci., U.S.A.*, 74, 5463-5467, (1977). The nucleotide and translated sequence for a cDNA coding for CEA-(g) is given hereinabove (TM-4 (CEA-(g))).

Example 4: Screening of a KG-1 cDNA Library with ³²P-labelled CEA Probe, LV7 (CEA-(A))

45 A segment of cDNA coding for a portion of carcinoembryonic antigen [LV7 or CEA-(a)] was radiolabelled by random priming and used to detect homologous sequences on filter replicas of a commercial cDNA library prepared from KG-1 cells in bacteriophage vector λ gt11 (Clontech Laboratories, Inc., Palo Alto, CA., U.S.A.). Hybridizations were performed at 68 °C in 2xSSPE (1xSSPE - 0.15 M NaCl, 0.01 M sodium phosphate and 1 mM EDTA, pH 7), 5x Denhardt's solution and 100 μg of denatured salmon sperm DNA per 50 ml, and post-hybridization washes were in 0.2xSSC, 0.25% sodium dodecyl sulfate.

Positive phage were picked, rescreened to homogeneity, and amplified for production of DNA. cDNA inserts were excised from phage DNA with EcoRI endonuclease and subcloned into the EcoRI site of the plasmid vector pBluescript KS. DNA sequencing on double-stranded DNA was by the method of Sanger et al, *supra*. The sequences of two different inserts from the KG-1 cDNA library are shown below:

55

pcKGCEAl:

1	acagcacagctgacagccgtactcaggaagcttctggatccttaggcttatctccacagag	60
5	51 gagaacacacaaggcagcagcagaccatggggccctctcagccctccctgcacacaccc MetGlyProLeuSerAlaProProCysThrHisLeu	120
10	121 atcaatttgaagggggctgctcacagcatcactttaaacttcttggaaatccggccaca IleThrTrpLysGlyValLeuLeuThrAlaSerLeuLeuAsnPheTrpAsnProProThr	180
15	181 actggccaaagtacgatttgaagcccagccacccaaagttctggggaaaggatgttctt ThrAlaGlnValThrIleGluAlaGlnProProLysValSerGluGlyLysAspValLeu	240
20	241 ctacttgtccacaatttgcggcagaatcttgcgttgcacattttgttacaaggccaaatg LeuLeuValHisAsnLeuProGlnAsnLeuAlaGlyTyrIleTrpTyrLysGlyGlnMet	300
25	301 acatacgttctaccattacattacatcatatgttagtagacggtcaaagaattatataatgg ThrTyrValTyrHisTyrIleThrSerTyrValValAspGlyGlnArgIleIleTyrGly	360
30	361 cctgcatacagtggaaagagaaagagtattccaatgcacccctgtgtccagaatgtc ProAlaTyrSerGlyArgGluArgValTyrSerAsnAlaSerLeuLeuIleGlnAsnVal	420
35	421 acgcaggaggatgcaggatcctacacccatcataaaggcgacgcgatggacttgg ThrGlnGluAspAlaGlySerTyrThrLeuHisIleIleLysArgArgAspGlyThrGly	480
40	481 ggagtaactggacatttcaccccttacacccatggagactccaaaggccctccatctcc GlyValThrGlyHisPheThrPheThrLeuHisLeuGluThrProLysProSerIleSer	540
45	541 agcagcaacttaatcccaggaggccatggaggctgtgtatcttaacctgtatccctgcg SerSerAsnLeuAsnProArgGluAlaMetGluAlaValIleLeuThrCysAspProAla	600
50	601 actccagccgcaagctaccagggtggatgaatggtcagaggccctccatgtactcacagg ThrProAlaAlaSerTyrGlnTrpTrpMetAsnGlyGlnSerLeuProMetThrHisArg	660
55	661 ttgcagctgtccaaaaccaacaggaccctttatattttgtgtcacaaggatattgca LeuGlnLeuSerLysThrAsnArgThrLeuPheIlePheGlyValThrLysTyrIleAla	720
60	721 ggaccctatgaatgtgaaatacggaaaccaggactgtggactgtccagccgcaggcacc GlyProTyrGluCysGluIleArgAsnProValSerAlaSerArgSerAspProValThr	780
65	781 ctgaatctccctccaaagctgtccaaaggccatcatacataatcaacaacttaaaccgg LeuAsnLeuLeuProLysLeuSerLysProTyrIleThrIleAsnAsnLeuAsnProArg	840
70	841 gagaataaggatgtcttaacccatgttgcacccatgttgcacactacacccat GluAsnLysAspValLeuThrPheThrCysGluProLysSerGluAsnTyrThrTyrIle	900
75	901 tggggctaaatggtcagaggccctccctgtcagtcggccagggtaaaggcgaccattggaaac TrpTrpLeuAsnGlyGlnSerLeuProValSerProArgValLysArgProIleGluAsn	960
80	961 aggatcctcatttccaaatgttcacgagaaatgaaacaggacccttatcaatgtgaaata ArgIleLeuIleLeuProAsnValThrArgAsnGluThrGlyProTyrGlnCysGluIle	1020
85	1021 cgggaccgatatggggcatccgcaggacttgcaccctgtatgtccatgttgc ArgAspArgTyrGlyGlyIleArgSerAspProValThrLeuAsnValLeuTyrGlyPro	1080

1081	gacctccccagcatttacccatttcattcacattaccgttcaggagaaaaacctctacttt AspLeuProSerIleTyrProSerPheThrTyrTyrArgSerGlyGluAsnLeuTyrPhe	1140
1141	tcctgcttcggtgagtctaaccacgggcacaatattctggacaattaatgggaagttt SerCysPheGlyGluSerAsnProArgAlaGlnTyrSerTrpThrIleAsnGlyLysPhe	1200
5		
1201	cagctatcaggacaaaagctctctatcccccaataactacaaggcatgtggctctat GlnLeuSerGlyGlnLysLeuSerIleProGlnIleThrThrLysHisSerGlyLeuTyr	1260
10		
1261	gcttgctctgttcgttaactcagccactggcaaggaaagctccaaatccatcacagtcaaa AlaCysSerValArgAsnSerAlaThrGlyLysGluSerSerLysSerIleThrValLys	1320
1321	gtctctgactggatattaccctgaattctactagttcctccaattccatttctccatg ValSerAspTrpIleLeuProEnd	1380
15		
141	gaatcacgaagagcaagagccactctgttccagaagccctataatctggagggtggacaac 1440	1440
141	tcgatgtaaatttcatggaaaacccctgtacactgacatgtgagccactcagaactcacc 1500	1500
1501	aaaatgttcgacaccataacaacagctactcaaactgtaaaccaggataagaatgtatg 1560	1560
1561	acttcacactgtggacagttttcaaagatgtcataacaagactcccatcatgacaagg 1620	1620
1621	ctccaccctctactgtctgctcatgcctgcctttcactggcaggataatgcgtcat 1680	1680
1681	tagaatttcacatgttagtagcttctgagggtaaacaacagactgtcagatgtcatctca 1740	1740
20		
1741	acctcaaactttacgttaacatctcaggaaatgtggctctccatcttgcatacagg 1800	1800
1801	ctcccaatagaaatgaaacacagagatattgtctgtgtttgcagagaagatgtttcta 1860	1860
1861	taaagagttaggaaagctgaaattatagtagactgtctctttaatgcacattgtgtggatg 1920	1920
1921	gctctcaccatttcctaagagatacagtgtaaaaaaacgtgacagtaatactgattctagca 1980	1980
1981	gaataaacatgttaccacatttgcaaaaaa 2010	2010
25		
	pCKGCEA2:	
1	gggtggatccctaggctcatctccataggggagaacacacatacagcagagaccatggga MetGly	59
30		
60	ccccctctcagccccctccctgcactcagcacatcaccttggaaagggtccctgctcacagca ProLeuSerAlaProProCysThrGlnHisIleThrTrpLysGlyLeuLeuThrAla 119	119
120	tcacttttaaacttctggAACCTGCCACCACTGCCAAGTAATAATTGAAGCCAGCCA SerLeuLeuAsnPheTrpAsnLeuProThrThrAlaGlnValIleIleGluAlaGlnPro 179	179
35		
180	cccaaagtctgaggggaaaggatgttcttacttgcaccaatttgcggccagaatctt ProLysValSerGluGlyLysAspValLeuLeuValHisAsnLeuProGlnAsnLeu 239	239
240	actggctacatctggtacaaggccaaatgacggacccatttaccattacatcatat ThrGlyTyrIleTrpTyrLysGlyGlnMetThrAspLeuTyrHisTyrIleThrSerTyr 299	299
40		
300	gtagtagacggtaaatttatatgggcctgcctacagtggacgagaaacagtatattcc ValValAspGlyGlnIleIleTyrGlyProAlaTyrSerGlyArgGluThrValTyrSer 359	359
360	aatgcattccctgtgtccagaatgtcacacaggaggatgcaggatcctacaccccttacac AsnAlaSerLeuLeuIleGlnAsnValThrGlnGluAspAlaGlySerTyrThrLeuHis 419	419
45		
420	atcataaaggcgaggcgatggactggaggagtaactggatatttcaactgtcaccttatac IleIleLysArgGlyAspGlyThrGlyGlyValThrGlyTyrPheThrValThrLeuTyr 479	479
480	tcggagactcccaaggcgctccatctccagcagcaacttaaaccacaggaggtcatggag SerGluThrProLysArgSerIleSerSerAsnLeuAsnProArgGluValMetGlu 539	539

540	gctgtgcgcttaatctgtgatcctgagactccggatgcaagctacacctgtggggctgaat AlaValArgLeuIleCysAspProGluThrProAspAlaSerTyrLeuTrpLeuLeuAsn	599
560	ggtcagaaccccttatgactcacaggttgcagctgtccaaaaccaacaggacccttat GlyGlnAsnLeuProMetThrHisArgLeuGlnLeuSerLysThrAsnArgThrLeuTyr	659
660	ctatttggtgtcacaaagtatattgcagggccctatgaatgtgaaatacggagggagtg LeuPheGlyValThrLysTyrIleAlaGlyProTyrGluCysGluIleArgArgGlyVal	719
720	agtgccagccgcagtgaccaggcacttgcacatctcccccgaagctgccatgccttac SerAlaSerArgSerAspProValThrLeuAsnLeuLeuProLysLeuProMetProTyr	779
780	atcaccatcaacaactaaacccagggagaagaaggatgttttagcccttacccgtgaa IleThrIleAsnAsnLeuAsnProArgGluLysLysAspValLeuAlaPheThrCysGlu	839
840	cctaagagtccgaaactacacacctacatttgggtggctaaatggtcagagcccccggcgt ProLysSerArgAsnTyrThrTyrIleTrpTrpLeuAsnGlyGlnSerLeuProValSer	899
900	ccgagggtaaagcgaccctattgaaaacaggataactcatttacccagggtgtcacgagaat ProArgValLysArgProIleGluAsnArgIleLeuIleLeuProSerValThrArgAsn	959
960	gaaacaggaccctatcaatgtgaaatacgggaccgatatggggatccgcgttaccc GluThrGlyProTyrGlnCysGluIleArgAspArgTyrGlyGlyIleArgSerAsnPro	1019
1020	gtcaccctgaatgtcctctatggtccagacacttcccccagaatttacccttacttccat ValThrLeuAsnValLeuTyrGlyProAspLeuProArgIleTyrProTyrPheThrTyr	1079
1080	taccgttcaggagaaaacctcgacttgtccgtttgcggactctaaccaccggcagag TyrArgSerGlyGluAsnLeuAspLeuSerCysPheAlaAspSerAsnProProAlaGlu	1139
1140	tattttggacaattaatgggaaagttcgtatcaggacaaaagctttatcccccaa TyrPheTrpThrIleAsnGlyLysPheGlnLeuSerGlyGlnLysLeuPheIleProGln	1199
1200	attactacaaatcatagccggctctatgtttgtttgttttttttttttttttttttttt IleThrThrAsnHisSerGlyLeuTyrAlaCysSerValArgAsnSerAlaThrGlyLys	1259
1260	gaaatctccaaatccatgatagtcaaaggctctgttttttttttttttttttttttttt GluIleSerLysSerMetIleValLysValSerGlyProCysHisGlyAsnGlnThrGlu	1319
1320	tctcattaaatggctgccacaatagagacacttgagaaaaagaacaggttgcata SerHisEnd	1379
1380	aaattcaagacaaaagaagaaaaaggctcaatgttattggactaaataatcaaaaggataa 1440	1439
1440	tgttttcataatttttatggaaaatgtgtctgatttttttttttttttttttttttttt 1500	1499
1500	tatgaacttttttcttcaagcaattggtaaagtatacttttttttttttttttttttt 1560	1559
1560	tttgcatttttttttttttttttttttttttttttttttttttttttttttttttttttt 1591	

It will be appreciated that the instant specification and claims are set forth by way of illustration and not limitation and that various modifications and changes may be made without departing from the scope of the present invention.

Claims

1. A nucleic acid comprising a base sequence which codes for a peptide sequence, characterized in that the group nucleic acid is a DNA selected from the following group of five sequences:

10

30

50

CAGCCGTGCTCGAACCGTTCCCTGGAGCCCCAAGCTCTCCTCCACAGGTGAAGACAGGGCCA

5

70

90

110

GCAGGAGACACCATGGGGCACCTCTCAGCCCCACTTCACAGAGTGCCTGTACCCCTGGCAG
MetGlyHisLeuSerAlaProLeuHisArgValArgValProTrpGln

10

130

150

170

GGGCTTCTGCTCACAGCCTCACTTCTAACCTTCTGGAACCCGCCACCACTGCCAGC
15 GlyLeuLeuLeuThrAlaSerLeuLeuThrPheTrpAsnProProThrThrAlaGlnLeu

190

210

230

20 ACTACTGAATCCATGCCATTCAATGTTGCAGAGGGGAAGGGAGGTCTTCTCCCTGTCCAC
ThrThrGluSerMetProPheAsnValAlaGluGlyLysGluValLeuLeuLeuValHis
25 250 270 290AATCTGCCAGCAACTTTTGCTACAGCTGGTACAAAGGGAAAGAGTGGATGCCAAC
25 AsnLeuProGlnGlnLeuPheGlyTyrSerTrpTyrLysGlyGluArgValAspGlyAsn

310

330

350

30 CGTCAAATTGTAGGATATGCAATTAGGAACCTCAACAAAGCTACCCAGGGCCGCAAACAGC
ArgGlnIleValGlyTyrAlaIleGlyThrGlnGlnAlaThrProGlyProAlaAsnSer

35

370

390

410

GGTCGAGAGACAATATACCCCAATGCATCCCTGCTGATCCAGAACGTCACCCAGAAATGAC
GlyArgGluThrIleTyrProAsnAlaSerLeuLeuIleGlnAsnValThrGlnAsnAsp

40

430

450

470

ACAGGATTCTACACCCCTACAAGTCATAAGTCAGATCTTGTGAATGAAGAAGCAACTGCA
ThrGlyPheTyrThrLeuGlnValIleLysSerAspLeuValAsnGluGluAlaThrGly

45

50

55

490 510 530
 CAGTTCCATGTATA
 5 GlnPheHisValTyrProGluLeuProLysProSerIleSerSerAsnAsnSerAsnPro
 550 570 590
 10 GTGGAGGACAGGATGCTGTGGCCTTCACCTGTGAACCTGAGACTCAGGACACAA
 ValGluAspLysAspAlaValAlaPheThrCysGluProGluThrGlnAspThrThrTyr
 610 630 650
 15 CTGTGGTGGATAAACAAATCAGAGCCTCCCCGGTCAGTCCAGGCTGCCAGGCTGCTGTCCAA
 LeuTrpTrpIleAsnAsnGlnSerLeuProValSerProArgLeuGlnLeuSerAsnGly
 20 670 690 710
 AACAGGACCCCTCACTCTACTCAGTGTACAAGGAATGACACAGGACCCATGAGTGTGAA
 AsnArgThrLeuThrLeuLeuSerValThrArgAsnAspThrGlyProTyrGluCysGlu
 25 730 750 770
 ATACAGAACCCAGTGAGTGCAGACCCAGTGACCCAGTCACCTGAAATGTCACCTATGCC
 IleGlnAsnProValSerAlaAsnArgSerAspProValThrLeuAsnValThrTyrGly
 30 790 810 830
 CCGGACACCCCCACCATTTCCCTTCAGACACCTATTACCGTCCAGGGGCAAA
 ProAspThrProThrIleSerProSerAspThrTyrTyrArgProGlyAlaAsnLeuSer
 35
 850 870 890
 CTCTCCTGCTATGCAGCCTCTAACCCACCTGCACAGTACTCCTGGCTTATCAATGGAA
 40 LeuSerCysTyrAlaAlaSerAsnProProAlaGlnTyrSerTrpLeuIleAsnGlyThr
 45 910 930 950
 TTCCAGCAAAGCACACAGAGCTTTATCCCTAACATCACTGTGAATAATAGTGGATCC
 PheGlnGlnSerThrGlnGluLeuPheIleProAsnIleThrValAsnAsnSerGlySer
 50 970 990 1010
 TATAACCTGCCACGCCAATAACTCAGTCACGGCTGCAACAGGACCA
 TyrThrCysHisAlaAsnAsnSerValThrGlyCysAsnArgThrThrValLysThrIle

1030 1050 1070
 5 ATAGTCACTGATAATGCTCTACCACAAGAAAATGGCTCTCACCTGGGCCATTGCAGGC
 IleValThrAspAsnAlaLeuProGlnGluAsnGlyLeuSerProGlyAlaIleAlaGly

 1090 1110 1130
 10 ATCTGATTGGAGTAGTGCCCCCTGGTCTGATAGCAGTAGCCCTGGCATGTTTCTG
 IleValIleGlyValValAlaLeuValAlaLeuIleAlaValAlaLeuAlaCysPheLeu

 1150 1170 1190
 15 CATTTCGGAGACCGGGCAGGGCAAGCGACCAGCGTGTCTCACAGAGCACAAACCTCA
 HisPheGlyLysThrGlyArgAlaSerAspGlnArgAspLeuThrGluHisLysProSer

 1210 1230 1250
 20 GTCTCCAAGCACACTCAGGACCCTCCAATGACCCACCTAACAGATGAATGAAGTTACT
 ValSerAsnHisThrGlnAspHisSerAsnAspProProAsnLysMetAsnGluValThr

 1270 1290 1310
 25 TATTCCTACCCCTGAACTTGAAGCCCAGCAACCCACACAACCAACTTCAGCCTCCCCATCC
 TyrSerThrLeuAsnPheGluAlaGlnGlnProThrGlnProThrSerAlaSerProSer

 1330 1350 1370
 30 CTAACAGCCACAGAAATAATTATTAGAAGTAAAAAGCAGTAATGAAACCTGTCTGC
 LeuThrAlaThrGluIleIleTyrSerGluValLysGln

 1390 1410 1430
 35 TCACTGCAGTGCTGATGTATTCAAGTCTCTCACCCCTCATCACTAGGAGATTCTTCCC

 1450 1470 1490
 40 CTGTAGGGTAGAGGGGTGGGACAGAAACAACTTCTCTACTCTTCTTAATAGGC

 1510 1530 1550
 45 ATCTCCAGGCTGCCCTGGTCACTGCCCTCTCTCAGTGTCAATTAGATGAAAGTACATTGGG

 1570 1590 1610
 50 AGTCTGTAGGAAACCCAACCTTCTTGTCAATTGAAATTGGCAAAGCTGACTTGGGAAAG

1630 1650 1670
 AGGGACCAGAACTTCCCTCCCTTCCCAACCTGGACTTGTAAACCTTGC
 5
 1690 1710 1730
 TGTCAGAGCACTCATTCCCTCCACCCCCAGTCCTGTCCTATCACTCTAAATTGGATTT
 10
 1750 1770 1790
 GCCATAGCCTTGAGGTTATGTCCTTTCCATTAAGTACATGTGCCAGGAAACAGCGAGAG
 15
 1810 1830 1850
 AGAGAAAGTAAACGGCAGTAATGCTTCTCCTATTCCTCCAAAGCCTTGTGTGAACCTAGCA
 20
 1870 1890 1910
 AAGAGAAAGAAATCAAAATATAACCAATAGTCAAATGCCACAGGTTGTCCACTGTCAG
 25
 1930 1950 1970
 GGTTGTCTACCTGTAGGATCAGGGTCTAAGCACCTGGTGCTTAGCTAGAAATACCACCTA
 30
 1990 2010 2030
 ATCCTTCTGGCAAGCCTGTCTTCAGAGAACCCACTAGAACCAACTAGGAAAAATCACTTG
 35
 2050 2070 2090
 CCAAAATCCAAGGCAATTCTGATGAAATGCAAAAGCACATATATGTTTAATATCTT
 40
 2110 2130 2150
 TATGGGCTCTGTTCAAGGCAGTGCTGAGAGGGAGGGTTATAGCTTCAGGAGGGAAACCA
 45
 2170 2190 2210
 CTTCTGATAAAAGACAAATCTGCTAGGAACCTGGAAAGGAATCAGAGAGCTGCCCTCAGC
 50

2230	2250	2270
GATTATTTAAATTGTTAAGAATACACAATTGGGGTATTGGGATTTCTCCTTTCTC		
5		
2290	2310	2330
TGAGACATTCGACCATTTAATTTTGTAACTGCTTATTTATGTGAAAAGGGTTATTTT		
10		
2350	2370	2390
ACTTAGCTTAGCTATGTCAGCCAATCCGATGCCCTAGGTGAAAGAAACCACCGAAATCC		
15		
2410	2430	2450
CTCAGGTCCCTGGTCAGGAGCCTCTCAAGATTTTTGTCAAGAGGCTCCAAATAGAAA		
20		
2470	2490	2510
ATAAGAAAAGGTTTCTTCATTCAATGGCTAGAGCTAGATTAACTCAGTTCTAGGCACC		
25		
2530	2550	2570
TCAGACCAATCATCAACTACCATCTATTCCATGTTGCACCTGTGCATTTCTGTTGC		
30		
2590	2610	2630
CCCCATTCACTTGTCAAGGAAACCTTGGCTCTGCTAAGGTGTATTTGGCTTGAGAAG		
35		
2650	2670	2690
TCGGAGCACCCCTACAGGGACACTATCACTCATGCTGGTGGCATTGTTACAGCTAGAAA		
40		
2710	2730	2750
CTGCACTGGTGCTAATGCCCTTGGAAATGGGGCTGTGAGGAGGAGGATTATAACTAG		
45		
2770	2790	2810
GCCTAGCCTCTTTAACAGCCTCTGAAATTATCTTTCTATGGGGCTATAAAATGT		
50		
2830	2850	2870
ATCTTATAATAAGGAAGGACAGGAGGAAGACAGGCAAAATGTACTTCTCACCCAGTCT		

2890 2910 2930
TCTACACAGATGGAATCTCTTGGGGCTAAGAGAAAGGTTTATTCTATATTGCTTACCT
5
2950 2970 2990
GATCTCATGTTAGGCCTAAGAGGCTTCCTCCAGGAGGATTAGCTTGGAGTTCTCTATACT
10
3010 3030 3050
CAGGTACCTCTTCAGGGTTTCTAACCCCTGACACGGACTGTGCATACTTTCCCTCATCC
15
3070 3090 3110
ATGCTGTGCTGTGTTATTTAATTTCTGGCTAAGATCATGTCTGAATTATGTATGAAA
20
3130 3150 3170
ATTATTCTATGTTTTATAATAAAATAATATATCAGACATCGAAAAAA,
25
30

35

40

45

50

55

(2)

5	10	30	50
CAGCCGTGCTCGAACGTTCTGGAGCCCCAAGCTCTCCTCCACAGGTGAAGACAGGGCCA			
10	70	90	110
GCAGGAGACACCATGGGGCACCTCTCAGCCCCACTTCACAGAGTGCCTGTACCCCTGGCAG MetGlyHisLeuSerAlaProLeuHisArgValArgValProTrpGln			
15	130	150	170
GGGCTTCTGCTCACAGCCTCACTTCTAACCTTCTGGAACCCGCCACCACTGCCAGCTC GlyLeuLeuLeuThrAlaSerLeuLeuThrPheTrpAsnProProThrThrAlaGlnLeu			
20	190	210	230
ACTACTGAATCCATGCCATTCAATGTTGCAGAGGGGAAGGAGGTTCTTCTCCTGTCCAC ThrThrGluSerMetProPheAsnValAlaGluGlyLysGluValLeuLeuLeuValHis			
25	250	270	290
AATCTGCCAGCAACTTTGGCTACAGCTGGTACAAAGGGAAAGAGTGGATGGCAAC AsnLeuProGlnGlnLeuPheGlyTyrSerTrpTyrLysGlyGluArgValAspGlyAsn			
310			
330			
350			
35	CGTCAAATTGTAGGATATGCAATAGGAACCTCAACAAGCTACCCAGGGCCCGCAAACAGC		
ArgGlnIleValGlyTyrAlaIleGlyThrGlnGlnAlaThrProGlyProAlaAsnSer			
370			
390			
410			
40	GGTCGAGAGACAATATACCCCAATGCATCCCTGCTGATCCAGAACGTCACCCAGAATGAC		
GlyArgGluThrIleTyrProAsnAlaSerLeuLeuIleGlnAsnValThrGlnAsnAsp			
45			
50			

430 450 470
 5 ACAGGATTCTACACCCCTACAAAGTCATAAAAGTCAGATCTTGTGAATGAAGAAGCAACTGG
 ThrGlyPheTyrThrLeuGlnValIleLysSerAspLeuValAsnGluGluAlaThrGly

 490 510 530
 10 CAGTTCCATGTATAACCGGAGCTGCCAAGGCCCTCCATCTCCAGCAACAACCTCCAACCC
 GlnPheHisValTyrProGluLeuProLysProSerIleSerSerAsnAsnSerAsnPro

 550 570 590
 15 GTGGAGGACAAGGATGCTGTGGCCTTCACCTGTGAACCTGAGACTCAGGACACAAACCTAC
 ValGluAspLysAspAlaValAlaPheThrCysGluProGluThrGlnAspThrThrTyr

 610 630 650
 20 CTGTGGTGGATAAACAAATCAGAGCCTCCGGTCAGTCCCAGGCTGCAGCTGTCCAATGGC
 LeuTrpTrpIleAsnAsnGlnSerLeuProValSerProArgLeuGlnLeuSerAsnGly

 25 670 690 710
 AACAGGACCCCTCACTCTACTCAGTGTACAAGGAATGACACAGGACCCATGAGTGTGAA
 AsnArgThrLeuThrLeuLeuSerValThrArgAsnAspThrGlyProTyrGluCysGlu

 30 730 750 770
 ATACAGAACCCAGTGAGTGCAGACCCAGTCACCTTGAAATGTCACCTATGGC
 IleGlnAsnProValSerAlaAsnArgSerAspProValThrLeuAsnValThrTyrGly

 35 790 810 830
 40 CCGGACACCCCCACCATTCCTTCAGACACCTATTACCGTCCAGGGGAAACCTCAGC
 ProAspThrProThrIleSerProSerAspThrTyrTyrArgProGlyAlaAsnLeuSer

 45

 50

 55

850

870

890

5 CTCTCCTGCTATGCAGCCTCTAACCCACCTGCACAGTACTCCTGGCTTATCAATGGAACA
 LeuSerCysTyrAlaAlaSerAsnProProAlaGlnTyrSerTrpLeuIleAsnGlyThr

910

930

950

10 TTCCAGCAAAGCACACAAGAGCTTTATCCCTAACATCACTGTGAATAATAGTGGATCC
 PheGlnGlnSerThrGlnGluLeuPheIleProAsnIleThrValAsnAsnSerGlySer

970

990

1010

15 TATACTGCCACGCCAATAACTCAGTCACTGGCTGCAACAGGACCACAGTCAGACGATC
 TyrThrCysHisAlaAsnAsnSerValThrGlyCysAsnArgThrThrValLysThrIle

1030

1050

1070

20 ATAGTCACTGAGCTAAGTCCAGTAGTAGCAAAGCCCCAAATCAAAGCCAGCAAGACCACA
 IleValThrGluLeuSerProValValAlaLysProGlnIleLysAlaSerLysThrThr

1090

1110

1130

25 GTCACAGGAGATAAGGACTCTGTGAACCTGACCTGCTCCACAAATGACACTGGAATCTCC
 ValThrGlyAspLysAspSerValAsnLeuThrCysSerThrAsnAspThrGlyIleSer

1150

1170

1190

30 ATCCGTTGGTTCTTCAAAAACCAGAGTCTCCCGTCCTCGGAGAGGATGAAGCTGTCCCAG
 IleArgTrpPhePheLysAsnGlnSerLeuProSerSerGluArgMetLysLeuSerGln

35

1210

1230

1250

40 GGCAACACCACCCCTCAGCATAAACCCCTGTCAAGAGGGAGGGATGCTGGGACGTATTGGTGT
 GlyAsnThrThrLeuSerIleAsnProValLysArgGluAspAlaGlyThrTyrTrpCys

45

50

55

1270 1290 1310
 5 GAGGTCTTCA-CCCAATCAGTAAGAACCAAGCGACCCATCATGCTGAACGTAAACTAT
 GluValPheAsnProIleSerLysAsnGlnSerAspProIleMetLeuAsnValAsnTyr

1330 1350 1370
 10 AATGCTCTACCACAAGAAAATGGCCTCTCACCTGGGCCATTGCTGGCATTGTGATTGGA
 AsnAlaLeuProGlnGluAsnGlyLeuSerProGlyAlaIleAlaGlyIleValIleGly

1390 1410 1430
 15 GTAGTGGCCCTGGTTGCTCTGATAGCAGTAGGCCCTGGCATGTTTCTGCATTCGGGAAG
 ValValAlaLeuValAlaLeuIleAlaValAlaLeuAlaCysPheLeuHisPheGlyLys

1450 1470 1490
 20 ACCGGCAGCTCAGGACCACTCCAATGACCCACCTAACAAAGATGAATGAAGTTACTTATTC
 ThrGlySerSerGlyProLeuGln

25 1510 1530 1550
 TACCCCTGAACTTGAAGGCCAGCAACCCACACAACCAACTTCAGCCTCCCCATCCCTAAC

30 1570 1590 1610
 AGCCACAGAAAATAATTATTAGAAGTAAAAAAGCAGTAATGAAACCTGAAAAA

35 1630
 AAAAAAAA

40

45

50

55

(3)

5

10

30

50

CAGCCGTGCTCGAACGCGTTCTGGAGCCCCAAGCTCTCCTCCACAGGGTGAACACAGGGCCA

10

70

90

110

GCAGGAGACACCATGGGGCACCTCTCAGCCCCACTTCACAGACTGGGTGACCCCTGGCAG
MetGlyHisLeuSerAlaProLeuHisArgValArgValProTrpGln

15

130

150

170

20 GGGCTTCTGCTCACAGCCTCACTTCTAACCTTCTGGAACCCGCCACCACTGCCAGCTC
GlyLeuLeuLeuThrAlaSerLeuLeuThrPheTrpAsnProProThrThrAlaGlnLeu

190

210

230

25 ACTACTGAATCCATGCCATTCAATGTTGGCTACAGCTGGTACAAAGGGAAAGGAGGTTCTCCTGTCCAC
ThrThrGluSerMetProPheAsnValAlaGluGlyLysGluValLeuLeuValHis

30

250

270

290

AATCTGCCAGCAACTTTTGCTACAGCTGGTACAAAGGGAAAGGAGTGGATGCCAAC
AsnLeuProGlnGlnLeuPheGlyTyrSerTrpTyrLysGlyGluArgValAspGlyAsn

35

310

330

350

CGTCAAATTGTAGGATATGCAATAGGAACCTAACAAAGCTACCCAGGGCCCGAAACAGC
ArgGlnIleValGlyTyrAlaIleGlyThrGlnGlnAlaThrProGlyProAlaAsnSer

40

370

390

410

GGTCGAGAGACAATATACCCCAATGCATCCCTGCTGATCCAGAACGTCACCCAGAAATGAC
GlyArgGluThrIleTyrProAsnAlaSerLeuLeuIleGlnAsnValThrGlnAsnAsp

45

430

450

470

50 ACAGGATTCTACACCCCTACAAAGTCATAAAAGTCAGATC'TTGTGAATGAACGAGCAAC'TGGA
ThrGlyPheTyrThrLeuGlnValIleLysSerAspLeuValAsnGluGluAlaThrGly

490 510 530
 5 CAGTTCCATGTATAACCGGAGCTGCCAAGCCCTCCATCTCCAGCAACAACTCCAACCT
 GlnPheHisValTyrProGluLeuProLysProSerIleSerSerAsnAsnSerAsnPro

 550 570 590
 10 GTGGAGGACAAGGATGCTGTGGCCTTCACCTGTGAAACCTGAGACTCAGGACACAAACCTAC
 ValGluAspLysAspAlaValAlaPheThrCysGluProGluThrGlnAspThrThrTyr

 610 630 650
 15 CTGTGGTGGATAACAATCAGAGCCTCCGGTCAGTCCCAGGCTGCAGCTGTCCAATGGC
 LeuTrpTrpIleAsnAsnGlnSerLeuProValSerProArgLeuGlnLeuSerAsnGly

 670 690 710
 20 AACAGGACCCCTCACTCTACTCAGTGTACAAGGAATGACACAGGACCCATGAGTGTGAA
 AsnArgThrLeuThrLeuLeuSerValThrArgAsnAspThrGlyProTyrGluCysGlu

 730 750 770
 25 ATACAGAACCCAGTGAGTGCAGACCCAGTGACCCAGTCACCTGAAATGTCACCTATGGC
 IleGlnAsnProValSerAlaAsnArgSerAspProValThrLeuAsnValThrTyrGly

 30
 790 810 830
 35 CCGGACACCCCCACCATTCCCTTCAGACACCTATTACCGTCCAGGGGAAACCTCAGC
 ProAspThrProThrIleSerProSerAspThrTyrArgProGlyAlaAsnLeuSer

 850 870 890
 40 CTCTCCTGCTATGCAGCCTCTAACCCACCTGCACAGTACTCCTGGCTTATCAATGGAAACA
 LeuSerCysTyrAlaAlaSerAsnProProAlaGlnTyrSerTrpLeuIleAsnGlyThr

 910 930 950
 45 TTCCAGCAAAGCACACAAGAGCTTTATCCCTAACATCACTGTGAATAATAGTGGATCC
 PheGlnGlnSerThrGlnGluLeuPheIleProAsnIleThrValAsnAsnSerGlySer

 50 970 990 1010
 55 TATACCTGCCACGCCAATAACTCAGTCACTCGCTGCAACAGGACCCACAGTCAGACGATC
 TyrThrCysHisAlaAsnAsnSerValThrGlyCysAsnArgThrThrValLysThrIle

1030

1050

1070

5 ATAGTCACTGATAATGCTCTACCACAAAGAAAATGGCCTCTCACCTGGGCCATTGCTGGC
 IleValThrAspAsnAlaLeuProGlnGluAsnGlyLeuSerProGlyAlaIleAlaGly

1090

1110

1130

10 ATTCTGATTGGACTAGTGGCCCTGGTTGCTCTGATAGCAGTAGCCCTGGCATGTTTCTG
 11 IleValIleGlyValValAlaLeuValAlaLeuIleAlaValAlaLeuAlaCysPheLeu

1150

1170

1190

15 CATTTCGGGAAGACCGGCAGCTCAGGACCACTCCAATGACCCACCTAACAGATGAATGA
 HisPheGlyLysThrGlySerSerGlyProLeuGln

1210

1230

1250

20 AGTTACTTATTCTACCCCTGAACTTGAAGCCCAGCAACCCACACAAACCAACTTCAGCCTC

1270

1290

1310

25 CCCATCCCTAACAGCCACAGAAATAATTATTAGAAGTAAAAAGCAGTAATGAAACCT

1330

30 GAAAAAAAAAAAAAAAAA

35

40

45

50

55

(4)

5	1 acagcacagctgacagccgtactcaggaagcttctggatcctaggcttatctccacagag	60
	61 gagaacacacaaggcagcagagaccatggggccctctcagccctccctgcacacaccc	120
	MetGlyProLeuSerAlaProProCysThrHisLeu	
10	121 atcacttggaaagggggtcctgctcacagcatactttaaacttctggaaatccgcccaca...	180...
	IleThrTrpLysGlyValLeuLeuThrAlaSerLeuLeuAsnPheTrpAsnProProThr	
15	181 actgcccagaatgcacgatgtggccagccaccctaaagttctgagggaaaggatgttctt	240
	ThrAlaGlnValThrIleGluAlaGlnProProLysValSerGluGlyLysAspValLeu	
20	241 ctacttgcacaaatttgcacaaatcttgcgttgcacattttgttacaaaggccaaatg	300
	LeuLeuValHisAsnLeuProGlnAsnLeuAlaGlyTyrIleTrpTyrLysGlyGlnMet	
25	301 acatacgtctaccattacattacatcatatgttagtagacggtaaagaattatataatggg	360
	ThrTyrValTyrHisTyrIleThrSerTyrValValAspGlyGlnArgIleIleTyrGly	
30	361 cctgcatacagtggaaagagaaagagtatattccatgcacccctgcgtgatccagaatgtc	420
	ProAlaTyrSerGlyArgGluArgValTyrSerAsnAlaSerLeuIleGlnAsnVal	
35	421 acgcaggaggatgcaggatccctacacccatcataaagcgacgcgttgcgtggactggaa	480
	ThrGlnGluAspAlaGlySerTyrThrLeuHisIleIleLysArgArgAspGlyThrGly	
40	481 ggagtaactggacatttcacccacccatggagactcccaagccctccatctcc	540
	GlyValThrGlyHisPheThrPheThrLeuHisLeuGluThrProLysProSerIleSer	
45	541 agcagcaacttaatccaggggccatggaggctgtgtatcttaacctgtgatcctgcg	600
	SerSerAsnLeuAsnProArgGluAlaMetGluAlaValIleLeuThrCysAspProAla	
50	601 actccagcccaagctaccagggtggatgaatggtcagagccctccatgtactcacagg	660
	ThrProAlaAlaSerTyrGlnTrpTrpMetAsnGlyGlnSerLeuProMetThrHisArg	
55	661 ttgcagctgtccaaaaccaacaggaccctttatattgggtgcacaaagtatattgca	720
	LeuGlnLeuSerLysThrAsnArgThrLeuPheIlePheGlyValThrLysTyrIleAla	
60	721 ggaccctatgaatgtgaaatacggaaaccaggatgtggccagccgcagtgcaccagg	780
	GlyProTyrGluCysGluIleArgAsnProValSerAlaSerArgSerAspProValThr	
65	781 ctgaatctccctccaaagctgtccaaaggccctacatcacaatcaacaacttaaaccgg	840
	LeuAsnLeuProLysLeuSerLysProTyrIleThrIleAsnAsnLeuAsnProArg	
70	841 gagaataaggatgtcttaacccatgtggacacttaagaggatggaaactacacccat	900
	GluAsnLysAspValLeuThrPheThrCysGluProLysSerGluAsnTyrThrTyrIle	
75	901 tggggctaaatggtcagggccctccgtcagtcggatggatggaaacggacccatgtggaaac	960
	TrpTrpLeuAsnGlyGlnSerLeuProValSerProArgValLysArgProIleGluAsn	
80	961 aggatccctatttacccatgtcacgagaaatgaaacaggacccatcaatgtgaaata	1020
	ArgIleLeuIleLeuProAsnValThrArgAsnGluThrGlyProTyrGlnCysGluIle	
85	1021 cgggaccgatatggtggcatccgcaggatgcaccaggatcaccctgttgcctctatggcca	1080
	ArgAspArgTyrGlyIleArgSerAspProValThrLeuAsnValLeuTyrGlyPro	

1081	gaccccccagcattacccttcatcacattaccgttcaggagaaaacctctacttt AspLeuProSerIleTyrProSerPheThrTyrTyrArgSerGlyGluAsnLeuTyrPhe	1140
1141	tcctgcttcggtagtctaaccacggcacaatattcttggacaattaatggaaagttt SerCysPheGlyGluSerAsnProArgAlaGlnTyrSerTrpThrIleAsnGlyLysPhe	1200
1201	cagctatcaggacaaaagctcttatcccccaaataactacaaagcatagtggcttat GlnLeuSerGlyGlnLysLeuSerIleProGlnIleThrThrLysHisSerGlyLeuTyr	1260
1261	gcttgctctgtcgtaactcagccactggcaagggaaagctccaaatccatcacagtcaaa AlaCysSerValArgAsnSerAlaThrGlyLysGluSerSerLysSerIleThrValLys	1320
1321	gtctctgactggatattaccctgaattctacttagttcctccatttcattttccatg ValSerAspTrpIleLeuProEnd	1380
1381	gaatcacgaagagcaagaccactctgttccagaagccctataatctggaggtggacaac	1440
1441	tcgatgtaaatttcatggaaaacccttgcacatgtgagccactcagaactcacc	1500
1501	aaaatgttcgacaccataacaacagctactcaaactgtaaaccaggataagaagtgtatg	1560
1561	acttcacactgtggacagttttcaaaatgtcataacaagactccccatcatgacaagg	1620
1621	ctccaccctctactgtctgcattgcctgcctttcacttggcaggataatgcagtcat	1680
1681	tagaatttcacatgttagtagttctggggtaacaacacagatgtcagatatgtcatttca	1740
1741	acctcaaacttttacatgtcatacatctcaggaaaatgtggctctccatcttcatacaggg	1800
1801	ctcccaatagaaaaacacacagatattgcctgtgtttcagagaagatgtttcttca	1860
1861	taaagagtaggaaagctgaaattatagtagagtctccattaaatgcacattgtgtgatg	1920
1921	gctctaccatttcttaagagatacagtgtaaaacgtgacagtaataactgttctagca	1980
1981	gaataaacatgttaccacatttgcaaaaaaa	2010

25

30

35

40

45

50

(5)

1	gggtggatcctaggctatctccatagggagaacacacatacagcagagaccatggg 5	59
MetGly		
60	ccccctcagccccccctgcactcagcacatcacctggaaaggggctcctgctcacagca ProLeuSerAlaProProCysThrGlnHisIleThrTrpLysGlyLeuLeuLeuThrAla	119
120	tcacttttaacttctggAACCTGCCACCCTGCCAAAGTAATAATTGAAGCCAGCCA 10 SerLeuLeuAsnPheTrpAsnLeuProThrThrAlaGlnValIleIleGluAlaGlnPro	179
180	cccaaagttctgagggaaaggatgttcttacttgcacattgtccccagaatctt ProLysValSerGluGlyLysAspValLeuLeuValHisAsnLeuProGlnAsnLeu	239
240	actggctacatctggtacaaaggcaatgacggacctctaccattacattacatcatat 15 ThrGlyTyrIleTrpTyrLysGlyGlnMetThrAspLeuTyrHisTyrIleThrSerTyr	299
300	gttagagacggtcaaatttatatatggcctgcctacagtggacgagaaacagtatattcc ValValAspGlyGlnIleIleTyrGlyProAlaTyrSerGlyArgGluThrValTyrSer	359
360	aatgcacatccctgctgatccagaatgtcacacaggaggatgcaggatcctacac 20 AsnAlaSerLeuLeuIleGlnAsnValThrGlnGluAspAlaGlySerTyrThrLeuHis	419
420	atcataaagcgaggcgatggactggaggagaactggatatttactgtcacctatac IleIleLysArgGlyAspGlyThrGlyGlyValThrGlyTyrPheThrValThrLeuTyr	479
480	tcggagactccaaagcgctccatccagcagcaacttaaaccaggaggatcatggag 25 SerGluThrProLysArgSerIleSerSerAsnLeuAsnProArgGluValMetGlu	539
540	gctgtgcgcttaatctgtgatccgtgactccggatgcaagctacccgtgggtgtgaat AlaValArgLeuIleCysAspProGluThrProAspAlaSerTyrLeuTrpLeuAsn	599
600	ggtcagaacccctccatgactcacaggttgcagctgtccaaaaccaacaggacccttat 30 GlyGlnAsnLeuProMetThrHisArgLeuGlnLeuSerLysThrAsnArgThrLeuTyr	659
660	ctatttgggtcacaaggatattgcagggccctatgaatgtgaataacggaggggagtg LeuPheGlyValThrLysTyrIleAlaGlyProTyrGluCysGluIleArgArgGlyVal	719
720	agtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 35 SerAlaSerArgSerAspProValThrLeuAsnLeuLeuProLysLeuProMetProTyr	779
780	atcaccatcaacaacttaaaccaggagaagaaggatgtgttagccttacccgtgaa IleThrIleAsnAsnLeuAsnProArgGluLysAspValLeuAlaPheThrCysGlu	839
840	cctaagagtgcggactcacacccatggggctaaatggtcagacgcctccggcgtc 40 ProLysSerArgAsnTyrThrTyrIleTrpTrpLeuAsnGlyGlnSerLeuProValSer	899
900	ccgagggtaaagcgaccattgaaaacaggataactcatttacccaggatgtc ProArgValLysArgProIleGluAsnArgIleLeuIleLeuProSerValThrArgAsn	959
960	gaaacaggaccctatcaatgtgaaatacgggaccgatatggtggcatccgcaga 45 GluThrGlyProTyrGlnCysGluIleArgAspArgTyrGlyGlyIleArgSerAsnPro	1019

1020 gtcaccctgaatgtccttatggccagaccccccagaatttacccttacttacccat 1079
 ValThrLeuAsnValLeuTyrGlyProAspLeuProArgIleTyrProTyrPheThrTyr
 5 1080 taccgttcaggagaaaacccgcacttgtcctgcggactctaaccaccggcagag 1139
 TyrArgSerGlyGluAsnLeuAspLeuSerCysPheAlaAspSerAsnProProAlaGlu
 1140 tattttggacaattaatgggaagttcagctatcaggacaaaagctttatccccaa 1199
 TyrPheTrpThrIleAsnGlyLysPheGlnLeuSerGlyGlnLysLeuPheIleProGln
 10 1200 attactacaatcatagcgggctctatgctgtcgtaactcagccactggcaag 1259
 IleThrThrAsnHisSerGlyLeuTyrAlaCysSerValArgAsnSerAlaThrGlyLys
 1260 gaaatctccaaatccatgatgtcaaagtctctggccatggaaaccagacagag 1319
 GluIleSerLysSerMetIleValLysValSerGlyProCysHisGlyAsnGlnThrGlu
 15 1320 tctcattaatggctgccacaatagagacactgagaaaaaagaacaggtgataccttcatg 1379
 SerHisEnd
 1380 aaatcaagacaaaagaagaaaaaggctcaatgttattggactaataatcaaaggataa 1439
 1440 tggtttcataatttttatggaaaatgtgctgattcttggatgtttattctccagatt 1499
 1500 tatgaactttttcttcagcaattggtaagtatactttgtaaacaaaaattgaaaca 1559
 1560 ttgcctttgccttatctgagtggccccc 1591

20

2. A replicable recombinant cloning vehicle having an insert comprising a nucleic acid of claim 1.

25 3. A cell that is transfected, infected or injected with a recombinant cloning vehicle of claim 2.

4. A method for preparing a polypeptide, said method comprising the steps of
 (a) culturing the cell of claim 3
 (b) recovering the polypeptide expressed by said cell.

30 5. A method for preparing an antibody directed against a polypeptide said method comprising the steps of
 (a) preparing said polypeptide by the method of claim 4
 (b) injecting said polypeptide into a host capable of producing antibodies and
 (c) recovering said antibodies.

35

Patentansprüche

1. Nucleinsäure, umfassend eine Basen-Sequenz, die für eine Peptid-Sequenz codiert, dadurch gekennzeichnet, daß die Gruppen-Nucleinsäure eine DNA ist, die aus der folgenden Gruppe von fünf Sequenzen ausgewählt ist:

40

45

50

55

10	30	50
CAGCCCGTGCCTGAAAGCGTTCTGGAGCCAAAGCTCTCCACAGGTGAAGACACGGCC		
5		
GCACGGAGACACCATGGGGCACCTCTCAGCCCCACTTCACAGAGTGCCTGTAACCCCTGGCAG		
MetGlyHisLeuSerAlaProLeuHisArgValArgValProTrpGln		
10		
GGCCTCTCTCACACCCCTCACTTCTAACCTCTGGAAACCCCCCACCACCTGCCAGCTC		
GlyLeuLeuLeuThrAlaSerLeuLeuThrPheAsnProPheThrAlaGlnLeu		
150	150	170
ACTACTGAATCCATGCCATTCATGTTGCAAGGGAGGGAGGAGGTCTCTCCCTGTCAC		
ThrThrGluSerMetPheAsnValAlaGluGlyLysGluValLeuLeuLeuValHis		
200	270	290
AATCTGCCCAAGCAAACCTTTCTGGCTACAGCTGGTACAAAGGGAAACAGCTGGATGCCAAC		
AsnLeuPheGlnGlnLeuPheGlyTyrSerThrPheTyrLysGlyGluArgValAspGlyAsn		
250		
310	330	350
CGTCAAATTCTAGGATATGCAAATAGCAACTCAACAAAGCTACCCCCAGGGCCAAACAGC		
ArgGlnIleValGlyTyrAlaIleGlyThrGlnGlnAlaThrPheGlyProAlaAsnSer		
370	390	410
GGTCGAGAGACAATATACCCCAATGCATCCCTGCTGATCCAGAACGTCACCCAGAACGAC		
GlyArgGluThrIleTyrProAsnAlaSerIleLeuIleGlnAsnValThrGlnAsnAsp		
430	450	470
ACAGGATTCTACACCCCTACAGTCATAAACGTCAGATCTTGTGAATGAAGAACGAACTGCA		
ThrGlyPheTyrThrLeuGlnValIleLysSerAspLeuValAsnGluAlaThrGly		
45		
50		

490 510 530
 CAGTTCCATGATACCCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAACCT
 5 GlnPheHisValTy:P:roGluLeu?roLysProSerileSerSerAsnAsnSerAsnPro

 550 570 590
 GTGGAGGACAAAGGATGCTGTCGCCCTTCACCTGTGAACCTGAGACTCAGGACACAAACCTAC
 10 ValGluAspLysAspAlaValAlaPheThrCysGluProGluThrGlnAspThrThrTyr

 610 630 650
 15 CTGTGGTGGATAAACAAATCAGAGCCTCCCGGTAGTCCCAGGCTGCAGCTGTCCAATGGC
 LeuTrpP:roIleAsnAsnGlnSerLeuProValSerProArgLeuGlnLeuSerAsnGly

 670 690 710
 20 AACAGGACCCCTCACTCTACTCAGTGTACAAGGAATGACACAGGACCCCTATCAGTGTGAA
 AsnAsnArgThrLeuThrLeuLeuSerValThrArgAsnAspThrGlyProTyrGluCysGlu

 730 750 770
 25 ATACAGAACCCCAAGTGAGTGCGAACCGCAGTGACCCAGTCACCTTGAATGTCACCTATGCC
 IleGlnAsnP:roValSerAlaAsnArgSerAspProValThrLeuAsnValThrTyrGly

 790 810 830
 30 CCGGACACCCCCACCCATTCCCTTCAGACACCTATTACCGTCCAGGGCAACCTCAGC
 ProAspThrP:roThrIleSerProSerAspThrTyrTyrArgProGlyAlaAsnLeuSer

 850 870 890
 35 CTCTCCTGCTATGCAGCCTCTAACCCACCTGCACAGTACTCCTGGCTTATCAAATGGAAACA
 LeuSerCysTyrAlaIleSerAsnProProAlaGlnTyrSerTrpIleAsnGlyThr

 910 930 950
 40 TTCCAGGAAAGCACAACAAGAGCTCTTATCCCTAACATCAGTGTGAAATPAGTGGATCC
 PheGlnGlnSerThrGlnGluLeuPheIleP:roAsnIleThrValAsnAsnSerGlySer

 970 990 1010
 45 TATAACCTGCCACGCCAATAACTCAGTCACUGCTGCAACAGGACCAACACTCAACACCGATC
 TyrThrCysHisAlaAsnAsnSerValThrGlyCysAsnArgThrThrValLysThrIle
 50

1030 1050 1070
 ATAGTCACTGATAATGCTCTACCAAGAAAATGGCCTCTCACCTGGGCCATTGCAGGC
 IleValThrAspAsnAlaLeuProGlnGluAsnGlyLeuSerProGlyAlaIleAlaGly
 5
 1090 1110 1130
 ATTGTGATTGGAGTAGTGGCCCTGGTTGCTCTGATACCAAGTAGCCCTGGCATGCTTTCTC
 IleValIleGlyValValAlaLeuValAlaLeuIleAlaValAlaLeuAlaCysPheLeu
 10
 1150 1170 1190
 CATTTCGGGAGGACCGGCAGGGCAAGCGACCACCGTGATCTCACAGAGCACAACCCCTCA
 15 HisPheGlyLysThrGlyArgAlaSerAspGlnArgAspLeuThrGluHisLysProSer
 1210 1230 1250
 CTCTCCAACCACACTCAGGACCACTCCAATGACCCACCTAACAAAGATGAATGAAGTTACT
 ValSerAsnHisThrGlnAspHisSerAsnAspProProAsnLysMetAsnGluValThr
 20
 1270 1290 1310
 TATTCCTACCCCTGAACCTTGAAGCCCAGCAACCCACACAAACCAACTTCAGCCTCCCCATCC
 TyrSerThrLeuAsnPheGluAlaGlnGlnProThrGlnProThrSerAlaSerProSer
 25
 1330 1350 1370
 CTAACAGCCACAGAAATAATTATTCAAGAAGTAAAAAAGCAGTAATGAAACCTGTCCTGC
 LeuThrAlaThrGluIleIleTyrSerGluValLysLysGln
 30
 1390 1410 1430
 TCACTGCAGTGCTGATGTATTCAGTCTCTCACCCCTCATCACTACGGAGATTCTTCCC
 35
 1450 1470 1490
 CTGTAGGGTAGAGGGGTGGGGACAGAAACAACTTTCTCTACTCTTCCCTTAATAGGC
 40
 1510 1530 1550
 ATCTCCAGGCTGCCCTGGTCAGTGCCCTCTCACGTGTCATAAGATGAAAGTACATTGGG
 45
 1570 1590 1610
 AGTCTGTAGGAAACCCAAACCTTCTGTCAATTGAAATTGGCAAAAGCTGACTTGGGAAAG
 50

1530

1650

1670

ACGGGACCAAGA=CTTCCCCCTCCCTTCCCCCTTTCCCAACCTGGACTTGTAAACTTCCC

5

1690

1710

1730

TGTTCAGAGCACTCATTCCTTCCCACCCCCAGTCCTGTCCTATCACTCTAAATTGGATTT

10

1750

1770

1790

GCCATAGCCCTGAGGTTATGTCCTTTCCATTAACTACATGTGCCAGGA=ACAGGGAGAG

15

1810

1830

1850

ACAGAAAGTAAACGGGAGTAAATGCTTCTCTATTCTCCAAAGCCTTGTTGTGAACTAGCA

20

1870

1890

1910

ACAGAAAGAAATCAAAATATAACCAATAAGTGAATGCCACAGGTTGTCCACTGTCAG

25

1930

1950

1970

GGTTGTCTACCTGTAGGATCAGGGTCTAAGCACCTTGGTCTTAGCTAGAAATACCAACCTA

30

1990

2010

2030

ATCCTTCTGGCAAGCCTGTCTCAGAGAACCCACTAGAACGAAACTAGGA=AAATCACTTG

35

2050

2070

2090

CCAAAATCCAAGGCAATTCCCTGATGGAAAATGCCAAAAGCACATATAATGTTTAAATACCTT

40

2110

2130

2150

TATGGGCTCTGTTCAAGGCAGTGCTGAGACGGAGGGGTTATAGCTTCAGGAGGGAAACCAAG

45

2170

2190

2210

CTTCTGATAAACAAATCTGCTAGGAACCTTGGAAAGGAATCAGAGAGCTGCCCTTCAGC

50

55

2230 2250 2270
 GATTATTTAATTGTTAACAAATACACAAATTGGGGTATTGGATTTTCTCCTTTCTC
 5
 2290 2310 2330
 TGAGACATTCACCACTTTAATTGGTAACTGCTTATTATGTCAGAAGGGTTATT
 10
 2350 2370 2390
 ACTTAGCTTAGCTATGTCAGCCAAATCCGATTGCCCTAGGTGAAAGAAACCAACCGAAATCC
 15
 2410 2430 2450
 CTCAGGTCCCTGGTCAGGAGCCTCTCAAGATTTTTGTCAAGAGCTCCAAATAGAAA
 20
 2470 2490 2510
 ATAAGAAAAAGCTTTCTTCATTCATGGCTAGAGCTAGATTTAACTCAGTTCTAGGCACC
 25
 2530 2550 2570
 TCAGACCAATCATCAACTACCATTCTATTCCATGTTGCACCTGTGCATTTCGTGGC
 30
 2590 2610 2630
 CCCCCATTCACTTTGTCAAGGAAACCTTGGCCTCTGCTAGGTGTATTCGGCCTTGAGAAG
 35
 2650 2670 2690
 TGGGAGCACCCCTACAGGGACACTATCACTCATGCTGGTGGCATGTTACAGCTAGAAG
 40
 2710 2730 2750
 CTGCACCTGGTGTAAATGCCCTGGAAATGGGGCTGTGAGGAGGAGGATTATAACTTAG
 45
 2770 2790 2810
 GCCTAGCCTCTTTAACAGCCCTCTGAAATTTATCTTTCTATGGGTCTATAAATGT
 50
 2830 2850 2870
 ATCTTATAATAAACGAAGGACAGGGAGGACAGGGAAATGTACTTCTCACCCACCT

2890

2910

2930

TCTACACAGATGAAATCTCTTGGGCTAAGAGAAAAGGTTTATTCTATATTGCTTACCT

5

2950

2970

2990

GATCTCATGTTAGGCCTAAGAGGCTTCTCCAGGAGGATTAGCTTGGACTCTCTATACT

10

3010

3030

3050

CAGGTACCTCTTCAGGGTTTCTAACCCCTGACACGGACTGTGCATACTTTCCCTCATCC

15

3070

3090

3110

ATGCTGTGCTGTGTTATTAAATTTCCTGGCTAAGATCATGTCATGAAATTATGTATGAAA

20

3130

3150

3170

ATTATTCTATGTTTATAAATAAAATATAATCAGACATCGAAAAAA,

25

30

35

40

45

50

55

(2)

5 10 30 50

CAGCCGTGCTCGAAGCGTTCTGGAGCCCAAGCTCTCCTCCACAGGTGAAGACAGGGCCA

10 70 90 110

GCAGGGAGACACCATGGGGCACCTCTCAGCCCCACTTCACAGAGTGCCTGTACCCCTGGCAG
MetGlyHisLeuSerAlaProLeuHisArgValArgValProTrpGln

15 130 150 170

GGGCTTCTGCTCACAGCCTCACTTCTAACCTTCTGGAACCCGCCACCACTGCCAGCTC
GlyLeuLeuLeuThrAlaSerLeuLeuThrPheTrpAsnProProThrThrAlaGlnLeu

20 190 210 230

ACTACTGAATCCATGCCATTCAATGTTGCAGAGGGAAAGGAGGTTCTTCTCTTGTCAC
ThrThrGluSerMetProPheAsnValAlaGluGlyLysGluValLeuLeuLeuValHis

25 250 270 290

AATCTGCCCAAGCAACTTTTGCTACAGCTGGTACAAAAGGGAAAGAGTGGATGGCAAC
AsnLeuProGlnLeuPheGlyTyrSerTrpTyrLysGlyGluArgValAspGlyAsn

30 310 330 350

CGTCAAATTGTAGGATATGCAATAGGAACCTCAACAAAGCTACCCAGGGCCCGCAACAGC
ArgGlnIleValGlyTyrAlaIleGlyThrGlnGlnAlaThrProGlyProAlaAsnSer

35 370 390 410

GGTCGAGAGACAAATATACCCCAATGCATCCCTGCTGATCCAGAACGTCACCCAGAAATGAC
GlyArgGluThrIleTyrProAsnAlaSerLeuLeuIleGlnAsnValThrGlnAsnAsp

45

50

430

450

470

ACAGGGATTCTACACCCCTACAGTCATAAGTCAGATCTTGTGAATGAAGAAGCAACTGG
 ThrGly?heTy:ThrLeuGlnValIleLysSerAspLeuValAsnGluGluAlaThrGly

5

490

510

530

CAGTTCCATGTATAACCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACCTCCAAACCT
 10 GlnPheHisValTyrProGluLeuProLysProSerIleSerSerAsnAsnSerAsnPro

550

570

590

15 GTGGAGGACAGGGATGCTGTGGCCTTCACCTGTGAACCTGAGACTCAGGACACAAACCTAC
 ValGluAspLysAspAlaValAlaPheThrCysGluProGluThrGlnAspThrThrTyr

20 CTGTGGTGGATAAACAAATCAGAGCCCTCCGGTCAGTCCCAGGCTGCAGCTGTCCAATGGC
 LeuTrpTrpIleAsnAsnGlnSerLeuProValSerProArgLeuGlnLeuSerAsnGly

670

690

710

25 AACAGGGACCCCTCACTCTACTCAGTGTACAAGGAATGACACAGGACCCATGAGTGTGAA
 AsnArgThrLeuThrLeuSerValThrArgAsnAspThrGlyProTyrGluCysGlu

30 730 750 770
 ATACAGAACCCAGTGAGTGGAAACCGCAGTGACCCAGTCACCTTGAATGTCACCTATGGC
 IleGlnAsnProValSerAlaAsnArgSerAspProValThrLeuAsnValThrTyrGly

35 790 810 830
 CCGGACACCCCCACCATTTCCCTTCAGACACCTATTACCGTCCAGGGGCAACCTCAGC
 ProAspThrProThrIleSerProSerAspThrTyrTyrArgProGlyAlaAsnLeuSer

40

45

50

55

550

870

890

CTCTCTGCTATGCCAGCCTCTAACCCACCTGCACAGTACTCCTGGCTTATCAAATGGAAACA
 LeuSerCystYrAlaAlaSerAsnProProAlaGlnTyrSerTrpLeuIleAsnGlyThr

5

910

930

950

TTCCAGGCAAGCACACAAAGAGCTCTTATCCCTAACATCACTGTGAATAATAAGTGGATCC
 PheGlnGlnSerThrGlnGluLeuPheIleProAsnIleThrValAsnAsnSerGlySer

10

970

990

1010

15

TATACCTGCCACGCCAATAACTCAGTCACTGGCTGCAACAGGACCACAGTCAAGACGATC
 TyrThrCysHisAlaAsnAsnSerValThrGlyCysAsnArgThrThrValLysThrIle

1030

1050

1070

20

ATAGTCACTGAGCTAAGTCCAGTAGTAGCAAAGCCCCAAATCAAAGCCAGCAAGACCACA
 IleValThrGluLeuSerProValValAlaLysProGlnIleLysAlaSerLysThrThr

1090

1110

1130

25

GTCACAGGGAGATAAGGACTCTGTGAAACCTGACCTGCTCCACAAATGACACTGGAAATCTCC
 ValThrGlyAspLysAspSerValAsnLeuThrCysSerThrAsnAspThrGlyIleSer

30

1150

1170

1190

ATCCGTTGGTTCTCAAAAACCAGAGTCTCCGTCTCGGAGAGGATGAAGCTGTCCCAG
 IleArgTrpPhePheLysAsnGlnSerLeuProSerSerGluArgMetLysLeuSerGln

35

1210

1230

1250

GGCAACACCACCTCAGCATAAACCTGTCAAGAGGGAGGATGCTGGGACGTATTGGTGT
 GlyAsnThrLeuSerIleAsnProValLysArgGluAspAlaGlyThrTyrTrpCys

40

45

50

55

1270

1290

1310

5 GAGGTCTTCAACCCAAATCAGTAAGAACCAAAAGCGACCCCATCATGCTGAACGTAAACTAT
GluValPheAsnProIleSerLysAsnGlnSerAspProIleMetLeuAsnValAsnTyr

1330

1350

1370

10 ATGCTCTACCAACAGGAAATGGCCTCTCACCTGGGGCCATTGCTGGCATGTGATTGGA
AsnAlaLeuProGlnGluAsnGlyLeuSerProGlyAlaIleAlaGlyIleValileGly

1390

1410

1430

15 GTAGTGGCCCTGGTGCTCTGATAGCAGTAGCCCTGGCATGTTTCTGCATTCGGGAAG
ValValAlaLeuValAlaLeuIleAlaValAlaLeuAlaCysPheLeuHisPheGlyLys

1450

1470

1490

20 ACCGGCAGCTCAGGACCACTCCAATGACCCACCTAACAAAGATGAATGAGTTACTTATTG
ThrGlySerSerGlyProLeuGln

1510

1530

1550

25 TACCCCTGAACCTTGTAGCCAGCAACCCACACAAACCAACTTCAGCCTCCCCATCCCTAAC

1570

1590

1610

30 AGCCACAGAAATTTATTTCAGAAGTAAAAAAAGCAGTAATGAAACCTGAAATTTAA

1630

35 ~~.....~~

40

45

50

55

(3)

5 10 30 50
 CAGCCGTCGCTCGAAGCGTTCTGGAGCCCAAGCTCTCCTCACAGGTGAAGCACAGGCCA
 10 70 90 110
 GCAGGAGACACCATGGGGCACCTCTCAGCCCCACTTCACAGAGTGCCTGTACCCCTGGCAG
 MetGlyHisLeuSerAlaProLeuHisArgValArgValProTrpGln
 15 130 150 170
 GGGCTTCGCTCACAGCCTCACTTCTAACCTTCTGGAACCCGCCACCACTCCCCAGCTC
 GlyLeuLeuLeuThrAlaSerLeuLeuThrPheTerPAsnProProThrThrAlaClnLeu
 20 190 210 230
 ACTACTGAATCCATGCCATTCAATGTTGCAGAGGGGAAGGAGGTTCTCTCCTTGCCAC
 ThrThrGluSerMetProPheAsnValAlaGluGlyLysGluValLeuLeuLeuValHis
 25 250 270 290
 AATCTGCCCAAGCAACTTTTGCTACAGCTGGTACAAAGGGAAAGAGCTGGATGCCAAC
 AsnLeuProGlnGlnLeuPheGlyTyrSerTrpTyrLysGlyGluArgValAspGlyAsn
 30 310 330 350
 CGTCAAATGTAGGATATGCAATAGGAACCTAACAGCTACCCCAAGGGCCCCAACAGC
 ArgGlnIleValGlyTyrAlaIleGlyThrGlnGlnAlaThrProGlyProAlaAsnSer
 35 370 390 410
 CGTCGAGAGACAAATACCCCAATGCATCCCTGCTGATCCAGAACGTCACCCAGAAATGAC
 GlyArgGluThrIleTyrProAsnAlaSerLeuLeuIleGlnAsnValThrGlnAsnAsp
 40 430 450 470
 ACAGGATTCTACACCCCTACAAAGTCATAAAAGTCAGATCTTGTGAATGAAAGCAAGCAGACTCGA
 ThrGlyPheTyrThrLeuGlnValIleLysSerAspLeuValAsnGluGluAlaThrGly

490 510 530
 5 CAGTTCCATGTATAACCCGGAGCTGCCCAAGCCCTCCATCTCCAGCAAACAACTCCAAACCC
 GlnPheHisValTy:P:ProGluLeuProLysProSerIleSerSerAsnAsnSerAsnPro
 550 570 590
 10 CTGGAGGACAGGGATGCTGTGGCCTTCACCTGTGAACCTGAGACTCAGGACACAAACCTAC
 ValGluAspLysAspAlaValAlaPheThrCysGluProGluThrGlnAspThrTh:Ty:
 610 630 650
 15 CTGTGGTGGATAAACAAATCAGAGCCTCCGGTCAGTCCCAGGCTGCAGCTGTCCAAATGGC
 LeuTrpTrpIleAsnAsnGlnSerLeuProValSerProArgLeuGlnLeuSerAsnGly
 670 690 710
 20 AACAGGACCCCTCACTCTACTCAGTGTCAAAAGGAATGACACAGGACCCCTATGACTGTGAA
 AsnArgTh:LeuThrLeuLeuSerValThrArgAsnAspThrGlyProTy:GluCysGlu
 730 750 770
 25 ATACAGAAACCCAGTGAGTGCGAACCGCAGTGACCCAGTCACCTTGAATGTCACCTATGGC
 IleGlnAsnProValSerAlaAsnArgSerAspProValThrLeuAsnValThrTy:Gly
 30
 790 810 830
 35 CCCGACACCCCCACCAATTCCCTTCAGACACCTATTACCGTCCAGGGGCAAACCTCAGC
 ProAspThr:ProThrIleSerProSerAspThrTy:Ty:ArgProGlyAlaAsnLeuSer...
 850 870 890
 40 CCTCTCTGCTATGCAGCCTCTAACCCACCTGCAACACTACTCCTGGCTTATCAATGGAAACA
 LeuSerCysTyrAlaAlaSerAsnProProAlaGlnTy:SerTrpLeuIleAsnGlyThr
 910 930 950
 45 TTCCAGCAACGCAACACAAAGACCTCTTATCCCTAACATCACTGTGAATAATAG:GGATCC
 PheGlnGlnSerTh:GlnGluLeuPheIleProAsnIleThrValAsnAsnSerGlySer
 970 990 1010
 50 TATACCTGCCACGCCAATAACTCAGTCACTCGCTGCAACAGGACCAAGTCAGACGGATC
 Ty:ThrCysHisAlaAsnAsnSerValThrGlyCysAsnArgThrThrValLysThrIle
 55

1030 1050 1070
5 ATAGTCACTGATAATGCTCTACCAACAAAGAAATGGCCTCTCACCTGGGGCCATTGCCTGGC
IleValThrAspAsnAlaLeuProGlnGluAsnGlyLeuSerProGlyAlaIleAlaGly
1090 1110 1130
10 ATGCTGATTCGACTAGTGGCCCTGGTTGCTCTGATAGCAGTACCCCTGGCATGTTTCTG
IleValIleGlyValValAlaLeuValAlaLeuIleAlaValAlaLeuAlaCysPheLeu
1150 1170 1190
15 CATTTCGGGAAAGACCCGGCAGCTCAGGACCACTCCAATGACCCACCTAACAAAGATGAATGA
HisPheGlyLysThrGlySerSerGlyProLeuGln
20 1210 1230 1250
20 AGTTACTTATTCTACCCCTGAACTTTGAAGCCCAGCAACCCACACAAACCAACTTCAGCCTC
25 1270 1290 1310
25 CCCATCCCTAACAGCCACAGAAATAATTATTAGAAAGTAAAAAAAGCAGTAATGAAACCT
30 1330
30 GAAAAA.....
35
40
45
50

(4)

5	! acagcacagctgacagccgtactcaggaagcttctggatcctaggcttatctccacacag	60
51	gagacacacaaacccagcagagaccatggggccctctcagccctcccccacacacccctc MetGlyProLeuSerAieProProCysThrHisIeu	120
121	atcacttggaaagggtccctgcacagcatcactttaaacttctggatccgcccaca. IleThrTrpLysGlyValLeuLeuThrAlaSerLeuLeuAsnPheTrpAsnProProTh:	180
181	actccccaaacccacccattgaaacccagccacccaaagttctgagggaaaggatcttctt ThrAlaGlnValThrileGluAlaGlnProProLysValSerGluGlyLysAspValLeu	240
241	ctacttgcacaaatttgcggcagaatcttgcgtggctacattggataaaaaggccaaatgg LeuLeuValHisAsnLeuProGlnAsnLeuAlaGlyTyrIleTrpTyrLysGlyGlnMet	300
301	acatacgtctaccattacattacatcatatgttagtagacggctaaagaattataatggg ThrTyrValTyrHisTyrIleThrSerTyrValValAspGlyGlnArgIleIleTyrGly	360
361	cctgcatacactgaaaggaaaggatattccatgcacccctgtgtatccaggatgtc ProAlaTyrSerGlyArgGluArgValTyrSerAsnAlaSerLeuLeuIleGlnAsnVal	420
421	acgcaggaggatgcaggatccacacccatataaagcgacgcgtggactgga ThrGlnGluAspAlaGlySerTyrThrLeuHisIleIleLysArgArgAspGlyThrGly	480
481	ggagtaactggacatttcacccacccataccacccatggagactccaaaggccatctcc GlyValThrGlyHisPheThrPheThrLeuHisLeuGluThrProLysProSerIleSer	540
541	acgcaggaggatgcaggatccacacccatggagactccaaaggccatctcc SerSerAsnLeuAsnProArgGluAlaMetGluAlaValIleLeuThrCysAspProAla	600
601	actccagccgcacgttaccacgtggatgtgtatgtcaggccctccatgtactcaccc ThrProAlaAlaSerTyrGlnTrpTrpMetAsnGlyGlnSerLeuProMetThrHisArg	660
661	ttgcacgtgtccaaaaacccacaggaccctttatattttgggtgtccacaaatattcc LeuGlnLeuSerLysThrAsnArgThrLeuPheIlePheGlyValThrLysTyrIleAla	720
721	ggcccccataatgtgaaatccggacccaggatgtgtccacccgtgtggatcc GlyProTyrGluCysGluIleArgAsnProValSerAlaSerArgSerAspProValThr	780
781	ctggatctcccccacagctgtccacccctacatcaacaaacttacccca LeuAsnLeuLeuProLysLeuSerLysProTyrIleThrIleAsnAsnLeuAsnProArg	840
841	gacaaatggatgtttacccatgtgttttttttttttttttttttttttttttt GluAsnLysAspValLeuThrPheThrCysGluProLysSerGluAsnTyThrTyrIle	900
901	tgggtggctaaatggtcagagccctccctgtcagtccagggtaaaggcaccattggaaaaac TrpPheLeuAsnGlyGlnSerLeuProValSerProArgValLysArgProIleGluAsn	960
961	aggatccatccatccaaatgttcacggagaaatgaaacaggacccatcaatgtgaaata ArgIleLeuIleLeuProAsnValThrArgAsnGluThrGlyProTyrGlnCysGluIle	1020
1021	cgggaccgtatggatgtggatccgcacgtgaccaggatccatgttttttttttttt ArgAspArgTyrGlyGlyIleArgSerAspProValThrLeuAsnValLeuTyrGlyPro	1080

1081	gacctccccggcatttacccatttcattcaccttattaccgttcaggagaaaaccttacttt AspLeuProSerIleTyrProSerPheThrTyrTyrArgSerGlyGluAsnLeuTyrPhe	1140
1141	tcctgcttcggtagtctaaccacgggcacaatattcttggacaattaatggaaagttt SerCysPheGlyGluSerAsnProArgAlaGlnTyrSerTrpThrIleAsnGlyLysPhe	1200
5		
1201	caacctatcaggcacaazagctcttatcccccaataactacaaaggcatatgggcttat GlnLeuSerGlyGlnIlysLeuSerIleProGlnIleThrThrLysHisSerGlyLeuTyr	1260
10		
1261	gcttgctctcttcgtactcagccactggcaaggaaagctccaaatccatcacagtczaa AlaCysSerValArgAsnSerAlaThrGlyLysGluSerSerLysSerIleThrValLys	1320
1321	gtctctgactggatattacccctgaattctacttagttccatccattttctccatg ValSerAspTrpIleLeuProEnd	1380
15		
1381	gaatcacgaagagcaagaccactctgttccagaagccctataatctggaggtggacaac	1440
1441	tcgatgtaaattcatggaaaaccctgtacctgacatgtgagccactcagaacctcacc	1500
1501	aaaatgttcgcacaccataacaacagctactcaaactgtaaaccaggataagaagtgtatg	1560
1561	acttcacactgtggacagttttcaaaagatgtcataacaagactccccatcatgacaagg	1620
1621	ctccaccctctactgtctcatgcctgcctttcaacttgcaggataatgcagtcat	1680
1681	tagaatttcacatgttagactgttgcagggtaaacac2cagatgtcagatgtcatctca	1740
1741	acctcaaactttacgtaaacatctcagggaaatgtggctctccatcttcatacaggg	1800
20		
1801	ctcccaatagaatgtacacagagatattgcctgtgtttcagagaaatgtgttctca	1860
1861	taazagagttagggaaagctgaaattataagttagactgtctccatgtgtggatg	1920
1921	gcttcaccattcctaagagatacagtgtaaacacgtgacagtaataactgattctagca	1980
1981	gzaataaacatgttaccacatttgcaaaaaaa	2010

25

30

35

40

45

50

(5)

1	gggtggatcctaggctatctccatagggagaacacacatacagcagagaccatgg MetGly	59
5		
50	ccccctctcagcgcctccctgcactcagcacatcacctggaaaggggctctgctcacagca ProLeuSerAlaProProCysThrGlnHisIleThrTrpLysGlyLeuLeuLeuThrAla	119
10	tcactttaaacttctggacacctgcccaccactgcccagaataattgaagcccagcca SerLeuLeuAsnPheTrpAsnLeuProThrThrAlaGlnValIleIleGluAlaGlnPro	179
15	cccaaagtctgagggaaaggatgttctacttgcacaaattgcccagaatctt ProLysValSerGluGlyLysAspValLeuLeuLeuValHisAsnLeuProGlnAsnLeu	239
20	actggctacatctggataaaaggcaatgacggaccttaccattacattacatcatat ThrGlyTyrIleTrpTyrLysGlyGlnMetThrAspLeuTyrHisTyrIleThrSerTyr	299
25	gtatgtacggtcaaattataatatgggcctgcctacagtggacgagaaacagtatattcc ValValAspGlyGlnIleIleTyrGlyProAlaTyrSerGlyArgGluThrValTyrSer	359
30	aatgcattccctgtatccagaatgtcacacaggaggatgcaggatcctacacccatcac AsnAlaSerLeuIleGlnAsnValThrGlnGluAspAlaGlySerTyrThrLeuHis	419
35	atcataaagcgaggcgatggactggaggactaactggatatttactgtcaccttatac IleIleLysArgGlyAspGlyThrGlyValThrGlyTyrPheThrValThrLeuTyr	479
40	tcggagactcczaagcgctccatctccagcagcaacttaaaccggagggtcatggag SerGluThrProLysArgSerIleSerSerAsnLeuAsnProArgGluValMetGlu	539
45	gctgtgcgttaatctgtatccctgagactccggatgcaagctacctgtgggtgtgaat AlaValArgLeuIleCysAspProGluThrProAspAlaSerTyrLeuTrpLeuLeuAsn	599
50	ggtcagaaccccttatgactcacaggttgcagctgtccaaaaccaacaggacccttat GlyGlnAsnLeuProMetThrHisArgLeuGlnLeuSerLysThrAsnArgThrLeuTyr	659
55	ctatttgttgcacaaagtataattgcaggccctatgaatgtgaaatacggaggggagt LeuPheGlyValThrLysTyrIleAlaGlyProTyrGluCysGluIleArgArgGlyVal	719
60	agtgcgcaggccgcgtgaccgcgtcacccctgaatctcccccgaagctgcctac SerAlaSerArgSerAspProValThrLeuAsnLeuLeuProMetProTyr	779
65	atcaccatcacaacttaaaccggagaagaaggatgtgttagccttacccatgt IleThrIleAsnAsnLeuAsnProArgGluLysAspValLeuAlaPheThrCysGlu	839
70	cctaagagtcggaaactacacccatgtggctaaatggctcagggcccccgtcgt ProLysSerArgAsnTyrThrTyrIleTrpTrpLeuAsnGlyGlnSerLeuProValSer	899
75	ccgagggtaaagcgacccattgaaaacaggataactcattctaccaggatgt ProArgValLysArgProIleGluAsnArgIleLeuIleLeuProSerValThrArgAsn	959
80	gaaacaggaccctatcaatgtgaaatacgggaccgatatggggcatccgcagtaaccca GluThrGlyProTyrGlnCysGluIleArgAspArgTyrGlyGlyIleArgSerAsnPro	1019

1020	gtcacccctgaaatccctctatggccagaccccccagaatttacccttacttcacccat ValThrLeuAsnValLeuTyrGlyProAspLeuProArgIleTyrProTyrPheThrTyr	1079
1080	taccgttcggcagaaaaacctcgacttgtcctgccttgcggactctaaccaccggcagag TyrArgSerGlyGluAsnLeuAspLeuSerCysPheAlaAspSerAsnProProAlaGlu	1139
5		
1140	tatttttggccatataatggaaagttcagctatcaggacaaaagctttatccccaa TyrPheTrpThrIleAsnGlyLysPheGlnLeuSerGlyGlnLysLeuPheIleProGln	1199
10		
1200	attactacaaatcatagcggctctatgcttgcgtactcaggccactggcaag IleThrThrAsnHisSerGlyLeuTyrAlaCysSerValArgAsnSerAlaThrGlyLys	1259
1260	gaaatctccatccatgatagtcaaagtctctggccctgcattggaaaccagacagag GluIleSerLysSerMetIleValLysValSerGlyProCysHisGlyAsnGlnThrGlu	1319
1320	tctcatataatggctgccacaatagagacactgagaaaaagaacaggttataccatcg SerHisEnd	1379
15		
1380	aaattcaagacaaagaagaaaaaggctcaatgttattggactaaataatcaaaggataa 1439	
1440	tgttttcataattttattggaaaatgtgctgatcttggatgtttattctccagatt 1499	
1500	tatgaactttttcttcagcaattggtaagttatactttgttaacaaaattgaaaca 1559	
1560	tttgcattttgtcttatctgagtgccccccc 1591	

20

2. Replizierbares rekombinantes Kloniergehikel mit einem eine Nucleinsäure nach Anspruch 1 umfassenden Insert.

25 3. Zelle, die mit einem rekombinanten Kloniergehikel nach Anspruch 2 transfiziert, infiziert oder injiziert ist.

4. Verfahren zur Herstellung eines Polypeptids, umfassend die Schritte
(a) des Kultivierens der Zelle nach Anspruch 3,
(b) des Gewinnens des durch diese Zelle exprimierten Polypeptids.

30 5. Verfahren zur Herstellung eines gegen ein Polypeptid gerichteten Antikörpers, umfassend die Schritte
(a) des Herstellens des Polypeptids durch das Verfahren des Anspruchs 4,
(b) des Injizierens des Polypeptids in einen Wirt, der zur Bildung von Antikörpern befähigt ist, und
(c) des Gewinnens der Antikörper.

35

Revendications

40 1. Acide nucléique comprenant une séquence de bases qui code pour une séquence peptidique, caractérisé en ce que le groupe d'acides nucléiques est de l'ADN choisi parmi le groupe de cinq séquences ci-après :

45

50

10

30

50

CAGCCCGTGCCTCGAAGCCGTTCTGGAGCCCCAAGCTCTCCTCCACAGGTGAAGACACGGCCA

6

70

90

110

GCACGGAGACACCATGGGGCACCTCTCAGCCCCACTTCACAGAGTGGGTGTACCCCTGGCAG
10 MetGlyHisLeuSerAlaProLeuHisArgValArgValProTrpGln

130

150

170

15 CGGCTTCTGCTCACAGCCCTCACTTCTAACCTTCTCGAACCCGGCCACCACTGCCAGCTC
GlyLeuLeuLeuThrAlaSerLeuLeuThrPheT:pAsnPheProThrThrAlaGlnLei

190

210

230

20 ACTACTGAAATCCATGCCATTCAATGTTGCACAGGGAAAGCAGGTTCTTCTCCTTGTCAT
ThrThrGluSerMetPheAsnValAlaGluGlyLysGluValLeuLeuLeuValHis
250 270 29025 ACTCTGCCCGAGCAACTTTTGCTACAGCTGGTACAAAGGGAAAGAGACTGGATGCCAAC
AsnLeuPheGlnGlnLeuPheGlyTyrSerTrpTyrLysGlyGluArgValAspGlyAsn

310

330

350

30 CGTCAAAATTGTAAGGATATGCCATTAGCAACTCAACAGCTACCCAGGGCCCCGCAAAACAGC
ArgGlnIleValGlyTyrAlaIleGlyThrGlnGlnAlaThrPheGlyProAlaAsnSer

35

370 390 410
GGTCGAGAGACATATAACCCCAATGCATCCCTGCTGATCCAGAACGTCACCCAGAAATGAC
GlyArgGluThrIleTyrProAsnAlaSerLeuLeuIleGlnAsnValThrGlnAsnAsp

40

430

450

470

ACAGGATTCTACACCCCTACAAGTCATAAGTCAGATCTTGTGAATGAAAGAAGCAACTGGA
ThrGlyPheTyrThrLeuGlnValIleLysSerAspLeuValAsnGluGluAlaThrGly

45

50

55

	490	510	530
5	CAGTTCCATGATACCCGGAGCTGCCAAGCCCTCCATCTCCAGCAACACTCCAAACCT GlnPheHisValTyrProGluLeuProLysProSerIleSerSerAsnAsnSerAsnPro		
	550	570	590
10	GTGGAGGACAAAGGATGCTGTGCCCTTCACCTGTGAAACCTGAGACTCAGGACACAAACCTAC ValGluAspLysAspAlaValAlaPheThrCysGluProGluThrGlnAspThrThrTyr		
	610	630	650
15	CTGTGGTGGATATAAAATCAGAGCCTCCCGGTCAAGTCCCAGGGCTGCAGCTGTCCAAATGGC LeuTrpPheAsnAsnGlnSerLeuProValSerProArgLeuGlnLeuSerAsnGly		
	670	690	710
20	AACAGGACCCCTCACTCTACTCAGTGTACAAGGAATGACACAGGACCCCTATGAGTGTCAA AsnArgThrLeuLeuSerValThrArgAsnAspThrGlyProTyrGluCysGlu		
	730	750	770
25	ATACAGAAACCCAGTGAGTGCGAACCGCAGTGACCCAGTCACCTTGAATGTCACCTATGCC IleGlnAsnProValSerAlaAsnArgSerAspProValThrLeuAsnValThrTyrGly		
	790	810	830
30	CCGGACACCCCCACCATTTCCCTTCAGACACCTATTACCGTCCAGGGGCAAACCTCAGC ProAspThrProThrIleSerProSerAspThrTyrTyrArgProGlyAlaAsnLeuSer		
	850	870	890
35	CTCTCCTGCTATGCCAGCCTCTAACCCACCTGCACAGTACTCCTGGCTTATCAATGGAAAC LeuSerCysteAlaAlaSerAsnProProAlaGlnTyrSerTrpLeuIleAsnGlyThr		
	910	930	950
40	TTCCAGGAAAGCACACAAAGAGCTCTTATCCCTAACATCAGTGTGAAATATAGTGGATCC PheGlnGlnSerThrGlnGluLeuPheIleProAsnIleThrValAsnAsnSerGlySer		
	970	990	1010
45	TATACCTGCCACCCCAATAACTCACTCACTGGCTGCCAACAGGACCAACAGTCAGAGACCATC TyrThrCysHisAlaAsnAsnSerValThrGlyCysAsnArgThrThrValLysThrIle		
50			
55			

1030 1050 1070
 5 ATAGTCACTGATATGCTCTACCAACAAATGGCTCTCACCTGGGCCATTGCCTGG
 IleValThrAspAsnAlaLeu?GlnGluAsnGlyLeuSerProGlyAlaIleAlaGly

 1090 1110 1130
 10 ATTTGATGGACTAGTGGCCCTGGCTCTGATACCAACTAGCCCTGGCATGTTTCTG
 IleValIleGlyValValAlaLeuValAlaLeuIleAlaValAlaLeuAlaCysPheLeu

 1150 1170 1190
 15 CATTTCGGAAAGACCGGGCAGGGCAAGGGACCCAGCGTGTCTCACAGAGCACAAACCTCA
 HisPheGlyLysThrGlyArgAlaSerAspGlnArgAspLeuThrGluHisLysProSer

 1210 1230 1250
 20 GTCTCCAAACACACTCAGGACCACTCCAATGACCCACCTAACAAAGATGATGAAAGTTACT
 ValSerAsnHisThrGlnAspHisSerAsnAspProProAsnLysMetAsnGluValThr

 1270 1290 1310
 25 TATTCCTACCCCTGAACTTTGAAGGCCAGCAACCCACACAACCAACTTCAGCCTCCCCATCC
 TyrSerThrLeuAsnPheGluAlaGlnGlnProThrGlnProThrSerAlaSerProSer

 1330 1350 1370
 30 CTAACAGCCACAGAAATAATTATTAGTAAAGTAAAAAAGCAGTAAATGAAACCTGTCCCTGC
 LeuThrAlaThrGluIleIleTyrSerGluValLysLysGln

 1390 1410 1430
 35 TCACCTGCAGTGCTGATGTTCAAGTCTCTCACCCCTCATCACTAGGAGATTCCTTCCC

 1450 1470 1490
 40 CTGTAGGGTAGAGGGGTGGGACAGAAACAACTTCTCTACTCTTCCCTAAATAGGC

 1510 1530 1550
 45 ATCTCCAGGCTGGCTGCTCACTGGCCCTCTCACGTGTCATACTGAAACTACATTGGC

 1570 1590 1610
 50 AGTCTGTAGGAAACCCAACCTTCTTGTCAATTGAAATTGGCAAAAGCTGACTTTGGCAAG

1630 1650 1670
 AGGGACCAGAACTTCCCCCTCCCTTCCCCAAACCTGGACTTGTAAACTTCC
 5
 1690 1710 1730
 TGTTCAAGAGCACTCATTCCCTCCACCCCCAGTCCTGTCCTATCACTCTAAATTGGATTT
 10
 1750 1770 1790
 GCCATAGCCTTGAGGTTATGTCCTTTCCATTAGTACATGTGCCAGGAAACACCCGAGAC
 15
 1810 1830 1850
 AGAGAAAGTAAACGGCAGTAAATGCTTCTCTATTCTCCAAAGCCTTGTGTAACTAGCA
 20
 1870 1890 1910
 AACAGAAAGAAATCAAAATATAACCAAATAGTCAAATGCCACACGGTTGTCCACTGTCAG
 25
 1930 1950 1970
 GGTGTCCTACCTGTAGGATCAGGCTCTAAGCACCTGGTGTCTAGCTAGAAATACCACTA
 30
 1990 2010 2030
 ATCCCTCTGGCAAGCCTGTCCTCAGAGAACCCACTAGAAGCAACTAGGAAAAATCACTTG
 35
 2050 2070 2090
 CCAAAATCCAAAGGCAATCCCTGATGCAAAATGCCAAACACATATATGTTAAATATCTT
 40
 2110 2130 2150
 TATGGGCTCTGTTCAAGGCAGTGCTGAGAGGGAGGGGTTATAGCTTCAGGAGGGAAACCAAG
 45
 2170 2190 2210
 CTTCTGATAAAGACAAATCTGCTAGGAACCTGGAAAGGAATCAGAGAGCTGCCCTTCAGC
 50

2230 2250 2270
 CATTATTTAAATTGTTAAAGAATACACAATTGGCGTATGGGATTTCTCCCTTTCTC
 5
 2290 2310 2330
 TGAGACATTCACCATTAACTTTGTAAGTGCCTATTTATGTAAGGGTTATTCT
 10
 2350 2370 2390
 ACTTAGCTTAGCTATGTCAGCCAATCCGATTGCCCTAGGTGAAAGAAACCAACCGAAATCC
 15
 2410 2430 2450
 CTCAGGTCCCTTGGTCAGGAGCCTCTCAAGATTTTGTCAAGGGCTCCAAATAGAAA
 20
 2470 2490 2510-
 ATAAAGAAAAGTTTCTTCATTCAATGGCTAGAGCTAGATTTAAGCTAGTTCTAGGCACC
 25
 2530 2550 2570
 TCAGACCAATCATCACTACCATTCTATTCCATGTTGCACCTGTGCATTTCTGTTGC
 30
 2590 2610 2630
 CCCCATTCACTTGTCAAGGAAACCTTGGCCTCTGCTAACGTGTATTTGGCCTTGAGAAG
 35
 2650 2670 2690
 TGGGAGCACCCCTACAGGGACACTATCACTCATGCTGGTGGCATTGTTACACCTAGAAG
 40
 2710 2730 2750
 CTGCACTGGTGTAAATGCCCTTGGAAATGGGCTGTGAGGAGGAGGATTATAACTTAG
 45
 2770 2790 2810
 CCCTAGCCCTTTTAAAGCCCTCTGAAATTTATCTTCTATGGGTCTATAATCT
 50
 2830 2850 2870
 ATCTTATAATAAAGCAAGGACACGGAGGAAAGACAGGGAAATGTAACCTCTCACCCACTCT

2890	2910	2930
TCTACACAGATGAAATCTCTTGGGGCTAAGACAAAAGGTTTATTCTATATTGCTTACCT		
5		
2950	2970	2990
GATCTCATGTTAGGCCTAAGAGCCCTTCTCCAGGAGGATTAGCTTGGAGTTCTCTATACT		
10		
3010	3030	3050
CAGGTACCTCTTCAGGGTTTCTAACCCCTGACACGGACTGTGCATACTTTCCCTCATCC		
15		
3070	3090	3110
ATGCTGTGCTGTGTTATTTAATTTTCTGGCTAAGATCATGTCCTGAATTATGTATGAAA		
20		
3130	3150	3170
ATTATTTCTATGTTTATAATAAAATAATATATCAGACATCGAAAAAA,		
25		
30		
35		
40		
45		
50		
55		

(2)

5 10 30 50

CAAGCCGTGCTCGAACGTTCTGGAGCCCCAAGCTCTCCACAGGTGAAGACAGGGCCA

10 20 90 110

GCAGGGAGACACCATGGGGCACCTCTCAGCCCCACTTCACAGAGTGCCTGTACCCCTGGCAG
MetGlyHisLeuSerAlaProLeuHisArgValArgValProTrpGln

15 130 150 170

GGGCTTCTGCTCACAGCCTCACTTCTAACCTTCTGGAACCCGCCACCACTGCCAGCTC
GlyLeuLeuLeuThrAlaSerLeuLeuThrPheTrpAsnProProThrThrAlaGlnLeu

20 190 210 230

ACTACTGAATCCATGCCATTCAATGTTGCAGAGGGGAAGGAGGTTCTTCTCCCTGTCCAC
ThrThrGluSerMetProPheAsnValAlaGluGlyLysGluValLeuLeuLeuValHis

25 250 270 290

AACTGCCCCAGCAACTTTTGCTACAGCTGGTACAAAGGGAAAGAGTGGATGGCAAC
AsnLeuProGlnGlnLeuPheGlyTyrSerTrpTyrLysGlyGluArgValAspGlyAsn

30 310 330 350

CGTCAAATTCTAGGATATGCAATAGGAACCTCAACAAAGCTACCCAGGGCCCCCAAACAGC
ArgGinIleValGlyTyrAlaIleGlyThrGlnGlnAlaThrProGlyProAlaAsnSer

35 370 390 410

GGTCGAGAGACAATATAACCCCAATGCCATCCCTGCTGATCCAGAACGTCACCCAGAAATGAC
GlyArgGluThrIleTyrProAsnAlaSerLeuLeuIleGlnAsnValThrGlnAsnAsp

45

50

55

430 450 470
 ACAGGATTCTACACCCCTACAGTCATAAAGTCAGATCTTGTGAATGAAAGAAGCAACTGGAA
 Th:GlyPheTyr:Th:LeuGlnValIleLysSerAspLeuValAsnGluGluAlaThrGly
 5
 490 510 530
 CAGTTCCATGATACCCGGAGCTGCCAAGCCCTCCATCTCCAGCAACAACTCCAAACCT
 10 GlnPheHisValTyr:?:GluLeuProLysProSerIleSerSerAsnAsnSerAsnPro
 550 570 590
 15 GTGGAGGACAGGATGCTGTGGCCTTCACCTGTGAACCTGAGACTCAGGACACAAACCTAC
 ValGluAspLysAspAlaValAlaPheThrCysGluProGluThrGlnAspThrThrTyr
 610 630 650
 20 CTGTGGTGGATAAACAAATCAGAGCCTCCGGTCAGTCAGTCCCAGGCTGCAGCTGTCCAATGGC
 LeuTrpTrpIleAsnAsnGlnSerLeuProValSerProArgLeuGlnLeuSerAsnGly
 670 690 710
 25 AACAGGACCCCTCACTCTACTCAGTGTACAAGGAATGACACAGGACCCATGAGTGTGAA
 AsnArgTh:LeuTh:LeuLeuSerValThrArgAsnAspThrGlyProTyrGluCysGlu
 30 730 750 770
 ATACAGAACCCAGTGAGTGCAGACCCAGTCACCTTGAAATGTCACCTATGGC
 IleGlnAsnProValSerAlaAsnArgSerAspProValThrLeuAsnValThrTyrGly
 35 790 810 830
 CCGGACACCCCCACCATTTCCCTTCAGACACCTATTACCGTCCAGGGGCAACCTCAGC
 ProAspThrProThrIleSerProSerAspThrTyrTyrArgProGlyAlaAsnLeuSer
 40
 45
 50
 55

550 670 890
 5 CTCTCTGCTATGCAGCCTCTAACCCACCTGCACAGTACTCCTGGCTTATCAATGGAAACA
 LeuSerCystYc:AlaAlaSerAsnPro?ProAlaGlnTyrSerTrpLeuIleAsnGlyThr

 910 930 950
 10 TTCCAGCAAGCACACAGAGCTCTTATCCCTAACATCACTGTGAATAATAGTGGATCC
 PheGlnGlnSerThrGlnGluLeuPheIleProAsnIleThrValAsnAsnSerGlySer

 970 990 1010
 15 TATACTGCCACGCCAATAACTCAGTCACTGGCTGCAACAGGACCACAGTCAAGACGATC
 TyrThrCysHisAlaAsnAsnSerValThrGlyCysAsnArgThrThrValLysThrIle

 1030 1050 1070
 20 ATAGTCACTGAGCTAAGTCCAGTAGTAGCAAAGCCCCAAATCAAAGCCAGCAAGACCACA
 IleValThrGluLeuSerProValValAlaLysProGlnIleLysAlaSerLysThrThr

 1090 1110 1130
 25 GTCACACGGAGATAAGGACTCTGTGAACCTGACCTGCTCCACAAATGACACTGGAAATCTCC
 ValThrGlyAspLysAspSerValAsnLeuThrCysSerThrAsnAspThrGlyIleSer

 1150 1170 1190
 30 ATCCGTTGGTTCTTCAAAAACCAGAGTCCTCCGTCTGGAGAGGGATGAAGCTGTCCCAG
 IleArgTrpPhePheLysAsnGlnSerLeuProSerSerGluArgMetLysLeuSerGln

 1210 1230 1250
 35 GGCACACCCACCCCTCAGCATAAACCCCTGTCAAGAGGGAGGATGCTGGGACGTATTGGTGT
 GlyAsnThrThrLeuSerIleAsnProValLysArgGluAspAlaGlyThrTyrTrpCys
 40

 45

 50

 55

1270

1290

1310

5 GAGGCTCTCAACCCAAATCAGTAAGAACCAAGCGACCCCATCATGCTGAACTGAAACTAT
 5 GluValPheAsnProIleSerLysAsnGlnSerAspProIleMetLeuAsnValAsnTyr

10 AATGCTCTACCACAAAGPAAAATGGCCTCTCACCTGGGGCCATTGCTGGCATTGTGATGGA
 10 AsnAlaLeuProGlnGluAsnGlyLeuSerProGlyAlaIleAlaGlyIleValIleGly

1390

1410

1430

15 CTAGTGGCCCTGGTTGCTCTGATAGCAGTAGGCCCTGGCATGTTTCTGCATTTCGGGAAG
 15 ValValAlaLeuValAlaLeuIleAlaValAlaLeuAlaCysPheLeuHisPheGlyLys

20 ACCGGCAGCTCAGGACCACCTCAATGACCCACCTAACAAAGATGAATGAAGTTACTTATTC
 20 ThrGlySerSerGly?soleuGln

25 TACCCCTGAACTTTGAAAGCCCAAGCAACCCACACAAACCAACTTCAGCCTCCCCATCCCTAAC

30 1570 1590 1610
 30 AGCCACAGAAATAATTATTTCAGAAGTAAAAAAAGCAGTAATGAAACCTGAAATAAAAAAA

35 1630
 35 AAAAAAA

40

45

50

55

(3)

5

10

30

50

CAGCCGTCCTCGAAAGCGTTCTGGAGCCCCAGCTCTCCACAGGTGAAACACACACGGCCA

10

70

90

110

GCAGGAGACACCATGGGGCACCTCTCAGCCCCACTTCACAGAGTGCCTGTAACCCCTGGCAG

15

Met:Gly:His:Leu:Ser:Ala:Pro:Leu:His:Arg:Val:Arg:Val:Pro:T:pGln

130

150

170

20 GGGCTTCGCTCACAGCCTCACTTCTAACCTTCTGGAACCCGCCACCACTGCCAGCTC
Gly:Leu:Leu:Leu:Thr:Ala:Ser:Leu:Leu:Thr:Pro:Leu:Trp:Asn:Pro:Pro:Thr:Thr:Ala:Gln:Leu

25

190

210

230

ACTACTGAATCCATGCCATTCAATGTTCCAGAGGGAAAGGACGTTCTTCTCCTTCTCCAC
Thr:Thr:Glu:Ser:Met:Pro:Pro:Leu:Asn:Val:Ala:Glu:Gly:Lys:Glu:Val:Leu:Leu:Val:His

30

250

270

290

AATCTGCCCCAGCAAATTGGCTACAGCTGGTACAAAGGGAAAGACTGGATGGCAAC
Asn:Leu:Pro:Gln:Gln:Leu:Pro:Leu:Gly:Ty:Ser:Trp:Tyr:Lys:Gly:Glu:Arg:Val:Asp:Gly:Asn

35

310

330

350

CGTCAAATTGGATATGCAATAGGAACCTCAACAAAGCTACCCAGGGCCCGCAACAGC
Arg:Gin:Leu:Val:Gly:Ty:Ala:Leu:Gly:Thr:Gln:Gln:Ala:Thr:Pro:Gly:Pro:Ala:Asn:Ser

40

370

390

410

GCTCGAGAGACAAATACCCCAATGGCATCCCTGCTGATCCAGAACGTCACCCAGAAATGAC
Gly:Arg:Glu:Thr:Leu:Tyr:Pro:Asn:Ala:Ser:Leu:Leu:Leu:Gln:Asn:Val:Thr:Gln:Asn:Asp

45

430

450

470

ACAGGGATTCTACACCCCTACAGTCATAAAAGTCAGATCTTGTGAATGAAACAGCAACTGGC
Thr:Gly:Pro:Leu:Gln:Val:Leu:Gly:Ser:Asp:Leu:Val:Asn:Glu:Glu:Ala:Thr:Gly

55

490 510 530
 CAGTTCCATGTATAACCGGAGCTGCCAAGCCCTCCATCTCCAGCAACAACCTCCAAACCT
 GlnPheHisValTyr:ProGluLeuProLysProSerIleSerSerAsnAsnSerAsnPro
 5
 550 570 590
 CTGGAGGACAAAGGATGCTGTGGCCTTCACCTGTGAACTGAGACTCAGGACACAAACCTAC
 10 ValGluAspLysAspAlaValAlaPheThrCysGluProGluThrGlnAspThrThrTyr
 610 630 650
 CTGTGGTGGATAAACAATCAGAGCCTCCGGTCAGTCCCAGGCTGCAGCTGTCCAATGGC
 LeuTrpTyr:LeuAsnAsnGlnSerLeuProValSerProArgLeuGlnLeuSerAsnGly
 670 690 710
 AACAGGACCCCTCACTCTACTCAGTGTCAAAAGGAATGACACAGGACCCATGACTGTGAA
 AsnArgThr:LeuThrLeuLeuSerValThrArgAsnAspThrGlyProTyr:GluCysGlu
 730 750 770
 ATACAGAACCCAGTGAGTGCGAACCGCAGTGACCCAGTCACCTTGAATGTCACCTATGGC
 IleGlnAsn?ProValSerAlaAsnArgSerAspProValThrLeuAsnValThrTyr:Gly
 790 810 830
 CCCGACACCCCCACCATTTCCCCCTCAGACACCTATTACCGTCCAGGGGCAAAACCTCAGC
 ProAspThr:ProThrIleSerProSerAspThrTyr:Arg:ProGlyAlaAsnLeuSer
 35
 850 870 890
 CTCTCCTGCTATGGAGCCCTCTAACCCACCTGCACAGTACTCCTGGCTTATCAAATGGAAAC
 40 LeuSerCysTyrAlaAlaSerAsnProProAlaGlnTyr:SerTrpLeuIleAsnGlyThr
 910 930 950
 TTCCAGCAARGCACACAAAGAGCTCTTATCCCTAACATCACTGTGAAATAATAGTGATCC
 PheGlnGlnSerThr:GlnGluLeuPheIleProAsnIleThrValAsnAsnSerGlySer
 45
 970 990 1010
 TATACCTGCCAACCCCAATAACTCAGTCACCTGGCTGCCAACAGGACCAACAGTCAAGACGATC
 TyrThrCysHisAlaAsnAsnSerValThrGlyCysAsnArgThrThrValIleAsnIle
 50
 55

1030 ATAGTCACTGATAATGCTCTACCAAGAAAATGGCCTCTCACCTGGGGCATTGCTGGC
 5 IleValThrAspAspAlaLeuProGlnGluAsnGlyLeuSerProGlyAlaIleAlaGly
 1050
 1090 1110 1130
 10 ATTGTGATTGGAGTACTGGCCCTGGTTGCTCTGATAGCAGTAGCCCTGCCATGTTTCTG
 15 IleValIleGlyValValAlaLeuValAlaLeuIleAlaValAlaLeuAlaCysPheLeu
 1150 1170 1190
 15 CATTTCGGGAAAGACCCGGCAGCTCAGGACCACTCCAATGCCAACCTAACAGATGAATGA
 HisPheGlyLysThrGlySerSerGlyProLeuGln
 20 1210 1230 1250
 20 AGTTACTTATTCTACCCCTGAACTTTGAGGCCAGCAACCCACACAAACCAACTTCAGCCCTC
 25 1270 1290 1310
 25 CCCATCCCTAACACCCACAGAAAATTTATTCAAGAAGTAAAAAGCAGTAATGAAACCT
 30 1330
 30 GAAAAA.....
 35
 40
 45
 50
 55

(4)

EP 0 346 710 B1

1081	gaccctccccacattacccttcattcacctattaccgttcaggagaaaacctctacttt	1140
	AspLeuProSerIleTyrProSerPhrThrTyrTyrArgSerGlyGluAsnLeuTyrPhe	
1141	tcctgccttcgggtggactctaaacccacggggcacaatattcttggacaattaatgggaagttt	1200
5	SerCysPheGlyGluSerAsnProArgAlaGlnTyrSerTrpThrIleAsnGlyLysPhe	
1201	cagctatcaaaaaacgcctctatccccaaataactacaagcatagtgggccttat	1260
	GlnLeuSerGlyGlnLysLeuSerIleProGlnIleThrThrLysHisSerGlyLeuTyr	
1261	gcttgcgttcgttcataactcagccacttggcaaggaaagctccaaatccatcacaatcaaa	1320
10	AlaCysSerValArgAsnSerAlaThrGlyLysGluSerSerLysSerIleThrValLys	
1321	gtctctgactggatattaccctgaattctacttagttcctccaattccatttctccatg	1380
	ValSerAspTerIleLeuProEnd	
1381	gaatcacgaagacaaagacccactctgttccagaagccctataatctggaggtggacaac	1440
1441	tcgtatgtaaatttcatggaaaacccttgcacatgtgagccactcagaactcacc	1500
1501	aaaatgttcgcacaccataacaacagctactcaaactgtataacaagactcccatcatgacaagg	1560
1561	acttcacactgtggacaggttttcaaaagatgtcataacaagactcccatcatgacaagg	1620
1621	ctccacccctctactgtctgtcatgcctgcctttcacttggcaggataatgcagtcat	1680
1681	tagaatttcacatgttagtagcttctgagggtaacaacacagagtgcagatatgtcatctca	1740
1741	acctcaaactttacgtAACATCTCAGGGAAATGTGGCTCTCCATTTGCTACAGGG	1800
20	1801 ctcccaatgtggaaatgtggaaacacagagatattgcctgtgtttgcagagaagatgggttcta	1860
1861	tazagagttaggaaagctgaaatttatagtagaggtctccctttaaatgcacattgtgtggatg	1920
1921	gctctaccatccataagagatacagtgtaaaaaacgtgcacagtaataactgattctagca	1980
1981	gaatcaaacatgttaccacatggcaaaaaaa	2010

25 end

30

35

40

45

50

55

(5)

1	gggtggatccaggctcatctccataggggagaacacacatacagcagagaccatggg MetGly	59
5		
50	ccctctcaccccctccctgactcagcacatcacctggaaaggggctccgtcacagca ProLeuSerAlaProProCysThrGlnHisIleThrTrpLysGlyLeuLeuLeuThrAla	119
10		
120	tcaactttaaaccttggAACCTGCCACCTGCCAAGTAATAATTGAAGCCAGCCA SerLeuLeuAsnPheTrpAsnLeuProThrThrAlaGlnValIleIleGluAlaGlnPro	179
15		
180	ccaaaggctggaggatgttctacttgcacaatttgcggcagaatctt ProLysValSerGluGlyLysAspValLeuLeuLeuValHisAsnLeuProGlnAsnLeu	239
20		
240	actggctacatctggcacaaaggccaaatgacggaccttaccattacattacatcatat ThrGlyTyrIleTrpTyrLysGlyGlnMetThrAspLeuTyrHisTyrIleThrSerTyr	299
25		
300	gtatgtacggctaaattatataatggggctgcctacagtggacgagaaacagtatattcc ValValAspGlyGlnIleIleTyrGlyProAlaTyrSerGlyArgGluThrValTyrSer	359
30		
360	aatgcatcccgtgtatccagaatgtcacacaggaggatgcaggatcctacaccc tacac AsnAlaSerLeuLeuIleGlnAsnValThrGlnGluAspAlaGlySerTyrThrLeuHis	419
35		
420	atcataaaggcaggcgatggactggaggagtaactggatatttactgtcaccttatac IleIleLysArgGlyAspGlyThrGlyValThrGlyTyrPheThrValThrLeuTyr	479
40		
480	tcggagactccaaaggcgctccatctccagcagcaacttaaaccggaggatcatggag SerGluThrProLysArgSerIleSerSerAsnLeuAsnProArgGluValMetGlu	539
45		
540	gctgtgcgttaatctgtgatccctgagactccggatgcaagctacctgtgggtctgaat AlaValArgLeuIleCysAspProGluThrProAspAlaSerTyrLeuTrpLeuLeuAsn	599
50		
600	ggtcagaacctccatgactcacaggttgcagctgtccaaaaccaacaggacccttat GlyGlnAsnLeuProMetThrHisArgLeuGlnLeuSerLysThrAsnArgThrLeuTyr	659
55		
660	ctatttgggtcacaaagtatattgcagggccctatgaatgtgaaatacggagggagtg LeuPheGlyValThrLysTyrIleAlaGlyProTyrGluCysGluIleArgArgGlyVal	719
60		
720	agtgccagccgcagtgaccctgacccactcacctgtaaatctcccccgaagctgcccattgccttac SerAlaSerArgSerAspProValThrLeuAsnLeuProLysLeuProMetProTyr	779
65		
780	atcaccatcaacaacttaaaccggaggagaagaaggatgtgttagccttacctgtgaa IleThrIleAsnAsnLeuAsnProArgGluLysAspValLeuAlaPheThrCysGlu	839
70		
840	cctaagagtccgaactcacctacatttggggctaaatggtcagagcctccggcagt ProLysSerArgAsnTyrThrTyrIleTrpTrpLeuAsnGlyGlnSerLeuProValSer	899
75		
900	ccgagggtaaagcgaccattgaaaacaggataactcatttacccactgtgtcacgagaaat ProArgValLysArgProIleGluAsnArgIleLeuLeuProSerValThrArgAsn	959
80		
960	gaaacaggaccctataatgtgaaatacgggaccgatatggtggcatccgcagtaaccca GluThrGlyProTyrGlnCysGluIleArgAspArgTyrGlyGlyIleArgSerAsnPro	1019
85		

1020	gtcacccctggatctccatggccagaccccccaataccctactcacctat ValThrLeuAsnValLeuTyrGlyProAspLeuProArgIleTyrProTyrPheThrTyr	1079
1080	taccgttcaggcggaaaacctcgacttgcctgcgttgcggactctaaccaccggcagag TyrArgSerGlyGluAsnLeuAspLeuSerCysPheAlaAspSerAsnProProAlaGlu	1139
5		
1140	tattttggacaattaatgggaagttcagctatcaggacaaaagctcttatccccaa TyrPheTrpThrIleAsnGlyLysPheGlnLeuSerGlyGlnLysLeuPheIleProGln	1199
10		
1200	attactacaaatcatagcgggctctatgtttgcgttgcgtactcggcaactggcaag IleThrThrAsnHisSerGlyLeuTyrAlaCysSerValArgAsnSerAlaThrGlyLys	1259
1250	gaaatctccaaatccatgtatgtcaaaatgtctctggccatggaaaccagacagag GluIleSerLysSerMetIleValLysValSerGlyProCysHisGlyAsnGlnThrGlu	1319
1320	tctcattaaatggctgccacaatagagacactgaaaaaaaagaacaggttataccttcatg SerHisEnd	1379
15		
1380	aaattcaagacaaaagaagaaaaaggctcaatgttattggactaaataatcaaaggataa	1439
1440	tgttttcataattttattggaaaatgtgtctgatcttggaaatgtttattctccagatt	1499
1500	tatgaacttttttcttcagcaattgttaagtgataactttgtaaacaaaaattgaaaca	1559
1560	tttgcctttgcctctatctgagtgccccccc 1591	

20

2. Véhicule de clonage recombinant apte à une réplication, comportant un produit d'insertion comprenant un acide nucléique selon la revendication 1.

25 3. Cellule qui a été transfectée, infectée par un véhicule de clonage recombinant selon la revendication 2, ou à laquelle on a injecté ce dernier.

4. Procédé pour préparer un polypeptide, ledit procédé comprenant les étapes consistant à :

30 (a) cultiver la cellule selon la revendication 3, et
(b) récupérer le polypeptide exprimé par ladite cellule.

5. Procédé pour préparer un anticorps dirigé contre un polypeptide, ledit procédé comprenant les étapes consistant à :

35 (a) préparer ledit polypeptide par le procédé selon la revendication 4,
(b) injecter ledit polypeptide dans un hôte capable de produire des anticorps, et
(c) récupérer lesdits anticorps.

40

45

50

55

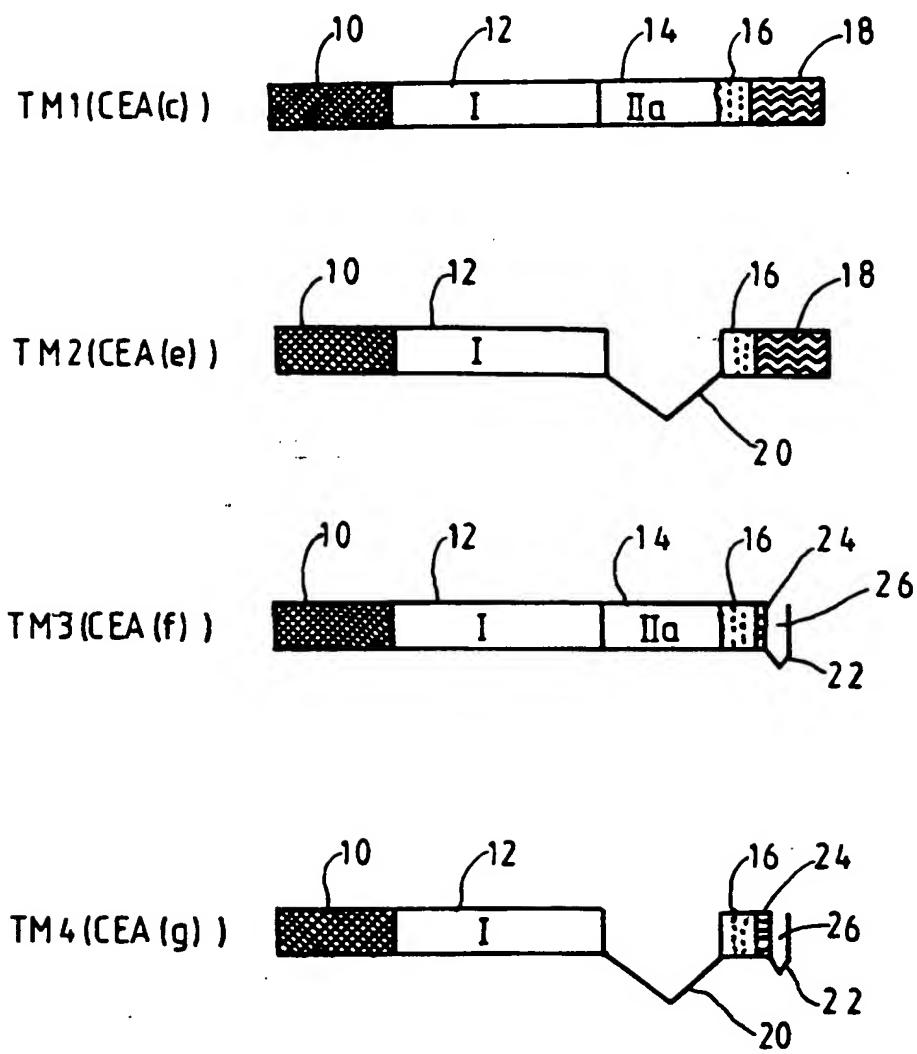


FIG.1